Problem 2:

(a) Prove that x is in the Cantor set iff x has a ternary expansion that uses only 0's and 2's.

(b) The Cantor-Lebesgue function is defined on the Cantor set by writing x's ternary expansion in 0's and 2's, switching 2's to 1's, and re-interpreting as a binary expansion. Show that this is well-defined and continuous, $F(0) = 0$, and $F(1) = 1$.

(c) Prove that F is surjective onto $[0, 1]$.

(d) Show how to extend F to a continuous function on $[0, 1]$.

Solution.

(a) The nth iteration of the Cantor set removes the open segment(s) consisting of all numbers with a 1 in the nth place of the ternary expansion. Thus, the numbers remaining after n iterations will have only 0's and 2's in the first n places. So the numbers remaining at the end are precisely those with only 0's and 2's in all places. (Note: Some numbers have a non-unique ternary representation, namely those that have a representation that terminates. For these, we choose the infinitely repeating representation instead; if it consists of all 0's and 2's, it is in the Cantor set. This works because we remove an *open* interval each time, and numbers with terminating representations are the endpoints of one of the intervals removed.)

(b) First, we show that this is well-defined. The only possible problem is that some numbers have more than one ternary representation. However, such numbers can have only one representation that consists of all 0's and 2's. This is because the only problems arise when one representation terminates and another doesn't. Now if a representation terminates, it must end in a 2 if it contains all 0's and 2's. But then the other representation ends with 12222... and therefore contains a 1.

Next we show F is continuous on the Cantor set; given $\epsilon > 0$, choose k such that $\frac{1}{2^k} < \epsilon$. Then if we let $\delta = \frac{1}{3^k}$, any numbers within δ will agree in their first k places, which means that the first k places of their images will also agree, so that their images are within $\frac{1}{2^k} < \epsilon$ of each other.

The equalities $F(0) = 0$ and $F(1) = 1$ are obvious; for the latter, $1 = 0.2222\ldots$ so $F(1) = 0.1111\cdots = 1$.

(c) Let $x \in [0, 1]$. Choose any binary expansion of x, replace the 0's with 2's, and re-interpret as a ternary expansion. By part (a), this will produce a member of the Cantor set whose image is x. (Note: Their may be more than one preimage of x, e.g. $F(\frac{1}{3}) = F(\frac{2}{3}) = \frac{1}{2}$.)

(d) First, note that F is increasing on the Cantor set C. Now let

$$G(x) = \sup\{F(y) : y \leq x, y \in C\}.$$

Note that $G(x) = F(x)$ for $x \in C$ because F is increasing. G is continuous at points not in C, because $\bar{C}$ is open, so if $z \in \bar{C}$, there is a neighborhood of z on which G is constant. To show that G is continuous on C, let $x \in C$ and use the continuity of F (part b) to

choose $\delta > 0$ such that $|G(x) - G(z) < \epsilon$ for $z \in C, |x-z| < \delta$. Choose $z_1 \in (x-\delta, x)$, $z_2 \in (x, x+\delta)$ and let $\delta' < \min(x - z_1, z_2 - x)$. Then for $|y - x| < \delta'$, if $y \in C$ we automatically have $|F(y) - F(x)| < \epsilon$. If $y \notin C$ but $y < x$, $G(x) > G(y) \geq G(z_1) > G(x) - \epsilon$; similarly, if $y \notin C$ but $y > x$, $G(x) < G(y) \leq G(z_2) < G(x) + \epsilon$.

□

Problem 3: Suppose that instead of removing the middle third of the segment at each step, we remove the middle ξ, where $0 < \xi < 1$.

(a) Prove that the complement of C_ξ is the union of open intervals with total length 1.

(b) Prove directly that $m_*(C_\xi) = 0$.

Solution.

(a) At the nth step (starting at $n = 0$), we remove 2^n segments, each of length $\xi\left(\frac{1-\xi}{2}\right)^n$. The total length of these segments is

$$\sum_{n=0}^{\infty} 2^n \xi \left(\frac{1-\xi}{2}\right)^n = \xi \sum_{n=0}^{\infty} \frac{1}{(1-\xi)^n} = \xi \frac{1}{1-(1-\xi)} = 1.$$

(b) If C_n is the set remaining after n iterations, then C_n is a union of 2^n segments of length $\left(\frac{1-\xi}{2}\right)^n$. So

$$m(C_n) = (1-\xi)^n.$$

Note that $m(C_n) \to 0$. Since each C_n is a covering of C by almost disjoint cubes, the infimum of the measures of such coverings is 0.

□

Problem 4: Construct a closed set $\hat{C}$ so that at the kth stage of the construction one removes 2^{k-1} centrally situated open intervals each of length ℓ_k, with

$$\ell_1 + 2\ell_2 + \cdots + 2^{k-1}\ell_k < 1.$$

(a) If ℓ_j are chosen small enough, then $\sum_{k=1}^{\infty} 2^{k-1}\ell_k < 1$. In this case, show that $m(\hat{C}) > 0$, and in fact,

$$m(\hat{C}) = 1 - \sum_{k=1}^{\infty} 2^{k-1}\ell_k.$$

(b) Show that if $x \in \hat{C}$, then there exists a sequence x_n such that $x_n \notin \hat{C}$, yet $x_n \to x$ and $x_n \in I_n$, where I_n is a sub-interval in the complement of $\hat{C}$ with $|I_n| \to 0$.

(c) Prove as a consequence that $\hat{C}$ is perfect, and contains no open interval.

(d) Show also that $\hat{C}$ is uncountable.

Solution.

(a) Let C_k denote the set remaining after k iterations of this process, with C_0 being the unit segment. Then

$$m([0,1] \setminus C_k) = \sum_{j=1}^{k} 2^j \ell_j$$

since $[0,1] \setminus C_k$ is a union of disjoint segments with this total length. Then

$$m(C_k) = 1 - \sum_{j=1}^{k} 2^j \ell_j.$$

Now $C_k \searrow \hat{C}$, so by Corollary 3.3,

$$m(\hat{C}) = \lim_{n\to\infty} m(C_k) = 1 - \sum_{k=1}^{\infty} 2^k \ell_k.$$

(b) For $k = 1, 2, \ldots$, let J_k be the interval of C_k which contains x. Let I_n be the interval in $\hat{C}^c$ which is concentric with J_{n-1}. (Thus, at the nth step of the iteration, the interval I_n is used to bisect the interval J_{n-1}.) Let x_n be the center of I_n. Then $x_n \in \hat{C}^c$. Moreover, $|x_n - x| \leq |J_{n-1}|$ since J_{n-1} contains both x_n and x. Since the maximum length of the intervals in C_n tends to 0, this implies $x_n \to x$. Finally, $x_n \in I_n \subset \hat{C}^c$, and $I_n \subset J_{n-1} \Rightarrow |I_n| \to 0$.

(c) Clearly $\hat{C}$ is closed since it is the intersection of the closed sets C_n. To prove it contains no isolated points, we use the same construction from the previous part. Let $x \in \hat{C}$. This time, let x_n be an endpoint of I_n, rather than the center. (We can actually take either endpoint, but for specificity, we'll take the one nearer to x.) Because I_n is constructed as an open interval, its endpoints lie in C_n. Moreover, successive iterations will not delete these endpoints because the kth iteration only deletes points from the interior of C_{k-1}. So $x_n \in \hat{C}$. We also have $|x_n - x| \leq J_n$ as before, so that $x_n \to x$. Hence x is not an isolated point. This proves that $\hat{C}$ is perfect.

(d) We will construct an injection from the set of infinite 0-1 sequences into $\hat{C}$. To do this, we number the sub-intervals of C_k in order from left to right. For example, C_2 contains four intervals, which we denote I_{00}, I_{01}, I_{10}, and I_{11}. Now, given a sequence $a = a_1, a_2, \ldots$ of 0's and 1's, let I_n^a denote the interval in C_n whose subscript matches the first n terms of a. (For instance, if $a = 0, 1, 0, 0, \ldots$ then $I_4^a = I_{0100} \subset C_4$.) Finally, let

$$x \in \bigcap_{n=1}^{\infty} I_n^a.$$

This intersection is nonempty because $I_{n+1}^a \subset I_n^a$, and the intersection of nested closed intervals is nonempty. On the other hand, it contains only one point, since the length of the intervals tends to 0. Thus, we have constructed a unique point in $\hat{C}$ corresponding to the sequence a. Since there is an injection from the uncountable set of 0-1 sequences into $\hat{C}$, $\hat{C}$ is also uncountable.

□

Problem 5: Suppose E is a given set, and and $\mathcal{O}_n$ is the open set

$$\mathcal{O}_n = \{x : d(x, E) < \frac{1}{n}\}.$$

Show that

(a) If E is compact, then $m(E) = \lim_{n\to\infty} m(\mathcal{O}_n)$.
(b) However, the conclusion in (a) may be false for E closed and unbounded, or for E open and bounded.

Proof. (a) First note that for any set E,

$$\bar{E} = \bigcap_{n=1}^{\infty} \mathcal{O}_n$$

since the closure of E consists of precisely those points whose distance to E is 0. Now if E is compact, it equals its closure, so

$$E = \bigcap_{n=1}^{\infty} \mathcal{O}_n.$$

Note also that $\mathcal{O}_{n+1} \subset \mathcal{O}_n$, so that $\mathcal{O}_n \searrow E$. Now since E is bounded, it is a subset of the sphere $B_N(0)$ for some N. Then $\mathcal{O}_1 \subset B_{N+1}(0)$, so that $m(\mathcal{O}_1) < \infty$. Thus, by part (ii) of Corollary 3.3,

$$m(E) = \lim_{n\to\infty} m(\mathcal{O}_n).$$

(b) Suppose $E = \mathbb{Z} \subset \mathbb{R}$. Then $m(\mathcal{O}_n) = \infty$ for all n, since $\mathcal{O}_n$ is a collection of infinitely many intervals of length $\frac{2}{n}$. However, $m(\mathbb{Z}) = 0$. This shows that a closed unbounded set may not work. To construct a bounded open counterexample, we need an open set whose boundary has positive measure. To accomplish this, we use one of the Cantor-like sets $\hat{C}$ from Problem 4, with the ℓ_j chosen such that $m(\hat{C}) > 0$. Let $E = [0,1]\backslash\hat{C}$. Then E is clearly open and bounded. The boundary of E is precisely $\hat{C}$, since $\hat{C}$ contains no interval and hence has empty interior. (This shows that the boundary of E contains $\hat{C}$; conversely, it cannot contain any points of E because E is open, so it is exactly equal to $\hat{C}$.) Hence $\bar{E} = E \cup \partial E = [0,1]$. Now

$$\mathcal{O}_n = \left\{x \in \mathbb{R} : d(x,E) < \frac{1}{n}\right\} = \left\{x \in \mathbb{R} : d(x,\bar{E}) < \frac{1}{n}\right\} = \left(-\frac{1}{n}, 1 + \frac{1}{n}\right).$$

Clearly $m(\mathcal{O}_n) \to 1$, but $m(E) = 1 - m(\hat{C}) < 1$.

□

Problem 6: Using translations and dilations, prove the following: Let B be a ball in $\mathbb{R}^d$ of radius r. Then $m(B) = v_d r^d$, where $v_d = m(B_1)$ and B_1 is the unit ball $\{x \in \mathbb{R}^d : |x| < 1\}$.

Solution. Let $\epsilon > 0$. Choose a covering $\{Q_j\}$ of B_1 with total volume less than $m(B_1) + \frac{\epsilon}{r^d}$; such a covering must exist because the $m(B_1)$ is the infimum of the volumes of such cubical coverings. When we apply the homothety $x \mapsto rx$ to $\mathbb{R}^d$, each Q_j is mapped to a cube Q_j' whose side length is r times the side length of Q_j. Now $\{Q_j'\}$ is a cubical covering of B_r with total volume less than $r^d m(B_1) + \epsilon$. This is true for any $\epsilon > 0$, so we must have $m(B_r) \le r^d m(B_1)$. Conversely, if $\{R_j\}$ is a cubical covering of B_r whose total volume is less than $m(B_r) + \epsilon$, we can apply the homothety $x \mapsto \frac{1}{r}x$ to get a cubical covering $\{R_j'\}$ of B_1 with total volume less than

$\frac{1}{r^d}(m(B_r)+\epsilon)$. This shows that $m(B_1) \le \frac{1}{r^d}(m(B_r)+\epsilon)$. Together, these inequalities show that $m(B_r) = r^d m(B_1)$. □

Problem 7: If $\delta = (\delta_1, \dots, \delta_d)$ is a d-tuple of positive numbers with $\delta_i > 0$, and $E \subset \mathbb{R}^d$, we define δE by

$$\delta E = \{(\delta_1 x_1, \dots, \delta_d x_d) : (x_1, \dots, x_d) \in E\}.$$

Prove that δE is measurable whenever E is measurable, and

$$m(\delta E) = \delta_1 \dots \delta_d m(E).$$

Solution. First we note that for an open set U, δU is also open. We could see this from the fact that $x \to \delta x$ is an invertible linear transformation, and therefore a homeomorphism. More directly, if $p \in U$, let $B_r(p)$ be a neighborhood of p which is contained in U; then if we define $\bar{\delta} = \min(\delta_1, \dots, \delta_d)$, we will have $B_{\bar{\delta}r}(\delta p) \subset \delta U$.

Next, we note that for any set S, $m_*(\delta S) = \delta_1 \dots \delta_d m_*(S)$. The proof of this is almost exactly the same as Problem 6: the dilation $x \mapsto \delta x$ and its inverse map rectangular coverings of S to rectangular coverings of δS and vice versa; but since the exterior measure of a rectangle is just its area (Page 12, Example 4), the infimum of the volume of rectangular coverings is the same as the infimum over cubical coverings. Hence a rectangular covering within ϵ of the infimum for one set is mapped to a rectangular covering within $\frac{\epsilon}{\delta_1 \dots \delta_n}$ for the other.

As a more detailed version of the preceding argument, suppose $\{Q_j\}$ is a cubical covering of S with $\sum |Q_j| < m_*(S)+\epsilon$. Then $\{\delta Q_j\}$ is a rectangular covering of δS with $\sum |\delta Q_j| < \delta_1 \dots \delta_d m_*(S) + \delta_1 \dots \delta_d \epsilon$. Now for each rectangle δQ_j we can find a cubical covering $\{Q'_{jk}\}$ with $\sum_k |Q'_{jk}| < |\delta Q_j| + \frac{\epsilon}{2^j}$. Then $\cap_{j,k} Q'_{jk}$ is a cubical covering of δS with $\sum_{j,k} |Q'_{jk}| < \delta_1 \dots \delta_d m_*(S) + (1+\delta_1 \dots \delta_d)\epsilon$. This implies that $m_*(\delta S) \le \delta_1 \dots \delta_d m_*(S)$. To get the reverse inequality we note that another δ-type transformation goes the other direction, i.e. $S = \delta'(\delta S)$ where $\delta' = (1/\delta_1, \dots, 1/\delta_d)$.

Now let $U \supset E$ be an open set with $m_*(U \setminus E) < \frac{\epsilon}{\delta_1 \dots \delta_d}$. Then $\delta U \supset \delta E$ is an open set. Moreover, $\delta(E \setminus U) = \delta E \setminus \delta U$, so $m_*(\delta U \setminus \delta E) = \delta_1 \dots \delta_d m_*(U \setminus E) < \epsilon$. Hence δE is also measurable. (Alternatively, we could prove that δE is measurable by appealing to Problem 8.) □

Problem 8: Suppose L is a linear transformation of $\mathbb{R}^d$. Show that if E is a measurable subset of $\mathbb{R}^d$, then so is $L(E)$, by proceeding as follows:

(a) Note that if E is compact, so is $L(E)$. Hence if E is an F_σ set, so is $L(E)$.

(b) Because L automatically satisfies the inequality

$$|L(x) - L(x')| \le M|x - x'|$$

for some M, we can see that L maps any cube of side length ℓ into a cube of side length $c_d M \ell$, with $c_d = 2\sqrt{d}$. Now if $m(E) = 0$, there is a collection of cubes $\{Q_j\}$ such that $E \subset \cap_j Q_j$, and $\sum_j m(Q_j) < \epsilon$. Thus $m_*(L(E)) \le c'\epsilon$, and hence $m(L(E)) = 0$. Finally, use Corollary 3.5. (Problem 4 of the next chapter shows that $m(L(E)) = |\det L| m(E)$.)

Solution.
(a) Since linear transformations on finite-dimensional spaces are always continuous, they map compact sets to compact sets. Hence, if E is compact, so is $L(E)$. Moreover, because $\mathbb{R}^d$ is σ-compact, any closed set is the countable union of compact sets. So if

$$E = \bigcap_{n=1}^{\infty} F_n$$

where F_n is closed, then for each n we have

$$F_n = \bigcap_{j=1}^{\infty} K_{nj}$$

where K_{nj} is compact; then

$$E = \bigcap_{j,n} K_{nj}$$

is a countable union of compact sets. Then

$$L(E) = \bigcap_{j,n} L(K_{nj})$$

is too, since $L(K_{nj})$ is compact. But compact sets are closed, so this shows that $L(E)$ is F_σ.

(b) Let x be a corner of a cube Q of side length ℓ. Then every point x' in the cube is a distance of at most $\sqrt{d}\ell$ away from x, since this is the distance to the diagonally opposite corner. Now $|x - x'| < \sqrt{d}\ell \Rightarrow |L(x) - L(x')| < \sqrt{d}M\ell$. Now if Q' is the cube of side length $2\sqrt{d}M\ell$ centered at x, the points on the exterior of the cube are all at least $\sqrt{d}M\ell$ away from x. $L(Q) \subset Q'$. Since a set of measure 0 has a cubical covering with volume less than ϵ, its image under L has a cubical covering with volume less than $2\sqrt{d}M\epsilon$. This implies that L maps sets of measure 0 to sets of measure 0.
Finally, let E be any measurable set. By Corollary 3.5, $E = C \cap N$ where C is an F_σ set and N has measure 0. We have just shown that $L(C)$ is also F_σ and $L(N)$ also has measure 0. Hence $L(E) = L(C) \cap L(N)$ is measurable.

□

Problem 9: Give an example of an open set $\mathcal{O}$ with the following property: the boundary of the closure of $\mathcal{O}$ has positive Lebesgue measure.

Solution. We will use one of the Cantor-like sets from Problem 4; let $\hat{C}$ be such a set with $m(hatC) > 0$. We will construct an open set whose closure has boundary $\hat{C}$. Let us number the intervals involved in the Cantor iteration as follows: If C_n is the set remaining after n iterations (with $C_0 = [0, 1]$), we number the 2^n intervals in C_n in binary order, but with 2's instead of 1's. For example, $C_2 = I_{00} \cap I_{02} \cap I_{20} \cap I_{22}$. The intervals in the complement of $\hat{C}$, denoted by subscripted J's, are named according to the intervals they bisected, by changing the last digit to a 1. For instance, in C_1, the interval J_1 is taken away to create the two intervals I_0 and I_2. In

the next iteration, I_0 is bisected by J_{01} to create I_{00} and I_{02}, while I_2 is bisected by J_{21} to create I_{20} and I_{21}, etc.

Having named the intervals, let $G = J_1 \cap J_{001} \cap J_{021} \cap J_{201} \cap J_{221} \cap \ldots$ be the union of the intervals in $\hat{C}^c$ which are removed during odd steps of the iteration, and $G' = [0,1] \setminus (G \cap \hat{C})$ be the union of the other intervals, i.e. the ones removed during even steps of the iteration. I claim that the closure of G is $G \cap \hat{C}$. Clearly this is a closed set (its complement in $[0,1]$ is the open set G') containing G, so we need only show that every point in $\hat{C}$ is a limit of points in G. To do this, we first note that with the intervals numbered as above, an interval $I_{abc\ldots}$ whose subscript is k digits long has length less than $\frac{1}{2^k}$. This is so because each iteration bisects all the existing I's. In addition, an interval $J_{abc\ldots}$ with a k-digit subscript has length less than $\frac{1}{2^{k-1}}$ because it is a subinterval of an I-interval with a $(k-1)$-digit subscript. Now let $x \in \hat{C}$. Then $x \in \cap_n C_n$ so for each n we can find an interval $I^{(n)}$ containing x which has an n-digit subscript. Let $J^{(n)}$ be the J-interval with an n-digit subscript, whose first $n-1$ digits match those of $I^{(n)}$. Then $I^{(n)}$ and $J^{(n)}$ are consecutive intervals in C_n. Since they both have length at most $\frac{1}{2^{n-1}}$, the distance between a point in one and a point in the other is at most $\frac{1}{2^{n-2}}$. Thus, if we let y_n be a sequence such that $y_n \in J^{(n)}$, then $y_n \to x$. Now let $y_{n'}$ be the subsequence taken for odd n, so that $y_{n'} \subset G$. Then we have constructed a sequence of points in G which converge to $x \in \hat{C}$.

We have shown that $\bar{G} = G \cap \hat{C}$. It only remains to show that $\partial(G \cap \hat{C}) = \hat{C}$. Clearly $\partial(G \cap \hat{C}) \subset \hat{C}$ since G is open and is therefore contained in the interior of $G \cap \hat{C}$. Now let $x \in \hat{C}$. By the same construction as above, we can choose a sequence $y_n \in J^{(n)}$ which converges to x. If we now take the subsequence $y_{\tilde{n}}$ over *even* n, then $y_{\tilde{n}} \in G'$ and $y_{\tilde{n}} \to x$. This proves that $x \in \partial(G \cap \hat{C})$. Hence we have shown that G is an open set whose closure has boundary $\hat{C}$, which has positive measure. □

Problem 11: Let A be the subset of $[0,1]$ which consists of all numbers which do not have the digit 4 appearing in their decimal expansion. Find $m(A)$.

Proof. A has measure 0, for the same reason as the Cantor set. We can construct A as an intersection of Cantor-like iterates. The first iterate is the unit interval; the second has a subinterval of length 1/10 deleted, with segments of lengths 3/10 and 6/10 remaining. (The deleted interval corresponds to all numbers with a 4 in the first decimal place.) The next has 9 subintervals of length 1/100 deleted, corresponding to numbers with a non-4 in the first decimal place and a 4 in the second. Continuing, we get closed sets C_n of length $(9/10)^n$, with $A = \cap C_n$. Clearly A is measurable since each C_n is; since $m(C_n) \to 0$, $m(A) = 0$. □

Problem 13:

(a) Show that a closed set is G_δ and an open set F_σ.
(b) Give an example of an F_σ which is not G_δ.
(c) Give an example of a Borel set which is neither G_δ nor F_σ.

Proof.

(a) Let U be open. As is well known, U is the union of the open rational balls that it contains. However, it is also the union of the *closed* rational balls that it contains. To prove this, let $x \in U$ and $r > 0$ such that $B_r(x) \subset U$. Choose a rational lattice point q with $|x - q| < \frac{r}{3}$, and a rational d with $\frac{r}{3} < d < \frac{r}{2}$. Then $\bar{B_d}(q) \subset B_r(x) \subset U$ and $x \in \bar{B_d}(q)$, so any $x \in U$ is contained in a closed rational ball within U. Thus, U is a union of closed rational balls, of which there are only countably many. For a closed set F, write the complement $\mathbb{R}^d \setminus F$ as a union of rational balls B_n; then $F = \cap B_n^c$ is a countable intersection of the open sets B_n^c, so F is G_δ.

(b) The rational numbers are F_σ since they are countable and single points are closed. However, the Baire category theorem implies that they are not G_δ. (Suppose they are, and let U_n be open dense sets with $\mathbb{Q} = \cap U_n$. Define $V_n = U_n \setminus \{r_n\}$, where r_n is the nth rational in some enumeration. Note that the V_n are also open and dense, but their intersection is the empty set, a contradiction.)

(c) Let $A = (\mathbb{Q} \cap (0,1)) \cup ((\mathbb{R} \setminus \mathbb{Q}) \cap [2,3])$ consist of the rationals in $(0,1)$ together with the irrationals in $[2,3]$. Suppose A is F_σ, say $A = \cup F_n$ where F_n is closed. Then

$$(\mathbb{R} \setminus \mathbb{Q}) \cap [2,3] = A \cap [2,3] = (\cup F_n) \cap [2,3] = \cup (F_n \cap [2,3])$$

is also F_σ since the intersection of the two closed sets F_n and $[2,3]$ is closed. But then

$$\mathbb{Q} \cap (2,3) = \cap (F_n^c \cap (2,3))$$

is G_δ because $F_n^c \cap (2,3)$ is the intersection of two open sets, and therefore open, for each n. But then if r_n is an enumeration of the rationals in $(2,3)$, $(F_n^c \cap (2,3)) \setminus \{r_n\}$ is also open, and is dense in $(2,3)$. Hence $\cap (F_n^c \cap (2,3)) \setminus \{r_j\}$ is dense in $(2,3)$ by the Baire Category Theorem. But this set is empty, a contradiction. Hence A cannot be F_σ.

Similarly, suppose A is G_δ, say $A = \cap G_n$ where G_n is open. Then

$$\mathbb{Q} \cap (0,1) = A \cap (0,1) = (\cap G_n) \cap (0,1) = \cap (G_n \cap (0,1))$$

is also G_δ since $G_n \cap (0,1)$ is the intersection of two open sets and therefore open. But then if $\{q_n\}$ is an erumeration of the rationals in $(0,1)$, $(G_n \cap (0,1)) \setminus \{q_n\}$ is open and is dense in $(0,1)$, so

$$\cap ((G_n \cap (0,1)) \setminus \{q_n\})$$

must be dense in $(0,1)$. But this set is empty, a contradiction. Hence A is not G_δ.

□

Problem 16: Borel-Cantelli Lemma: Suppose $\{E_k\}$ is a countable family of measurable subsets of $\mathbb{R}^d$ and that

$$\sum_{k=1}^{\infty} m(E_k) < \infty.$$

Let

$$E = \{x \in \mathbb{R}^d : x \in E_k \text{ for infinitely many } k\} = \limsup E_k.$$

- Show that E is measurable.
- Prove $m(E) = 0$.

Solution.

- Let

$$B_n = \bigcup_{k=n}^{\infty} E_k$$

be the set of x which are in some E_k with $k \geq n$. Then x is in infinitely many E_k iff $x \in B_n$ for all n, so

$$E = \bigcap_{n=1}^{\infty} B_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} E_k.$$

This is a countable intersection of a countable union of measurable sets, and hence is measurable.

- Let $\epsilon > 0$. Since $\sum m(E_k)$ converges, $\exists N$ such that

$$\sum_{k=N}^{\infty} m(E_k) < \epsilon.$$

Then

$$m(B_N) = m\left(\bigcup_{k=N}^{\infty} E_k\right) \leq \sum_{k=N}^{\infty} m(E_k) < \epsilon$$

by subadditivity. But $m(\cap B_n) \leq m(B_N)$ by monotonicity, so $m(E) < \epsilon$ for all ϵ. Hence $m(E) = 0$.

□

Problem 17: Let $\{f_n\}$ be a sequence of measurable functions on $[0, 1]$ with $|f_n(x)| < \infty$ for a.e. x. Show that there exists a sequence c_n of positive real numbers such that

$$\frac{f_n(x)}{c_n} \to 0 \quad \text{a.e. } x.$$

Solution. We are given that for each n,

$$m\left(\bigcap_{k=1}^{\infty} \left\{x : |f_n(x)| > \frac{k}{n}\right\}\right) = 0$$

since this set is precisely the set where $|f(x)| = \infty$. Since these sets are nested, this implies

$$\lim_{k\to\infty} m\left(\left\{x : |f_n(x)| > \frac{k}{n}\right\}\right) = 0.$$

Hence, $\exists c_n$ such that

$$m\left(\left\{x : |f_n(x)| > \frac{c_n}{n}\right\}\right) < \frac{1}{2^n}.$$

Define

$$E_n = \left\{x : |f_n(x)| > \frac{c_n}{n}\right\}.$$

Then $m(E_n) < \frac{1}{2^n}$, so

$$m\left(\bigcap_{m=1}^{\infty}\bigcup_{j=m}^{\infty} E_j\right) = 0$$

by the Borel-Cantelli lemma. But the complement of this set consists of precisely those points that are in finitely many E_n, i.e. those points for which $\frac{f_n(x)}{c_n}$ is eventually less than $\frac{1}{2^n}$. Hence we have found a set of measure 0 such that $\frac{f_n(x)}{c_n} \to 0$ on the complement. □

Problem 18: Prove the following assertion: Every measurable function is the limit a.e. of a sequence of continuous functions.

Proof. Let $f : \mathbb{R} \to \mathbb{R}$ be measurable. (The problem didn't specify whether f can have $\pm\infty$ as a value, but I'm assuming not.) Let $B_n = [-n, n]$. Then by Lusin's Theorem, there exists a closed (hence compact) subset $E_n \subset B_n$ with $m(B_n \backslash E_n) < \frac{1}{2^n}$ and f continuous on E_n. Then by Tietze's Extension Theorem, we can extend f to a continuous function f_n on all of $\mathbb{R}$, where $f_n = f$ on E_n. (Explicitly, such an extension could work as follows: Define $f_n : \mathbb{R} \to \mathbb{R}$ by $f_n(x) = f(x)$ for $x \in E_n$; for $x \notin E_n$, since the complement is open, x is in some open interval $(a, b) \subset E_n^c$ or in some unbounded open interval $(-\infty, a) \subset E_n^c$ or $(b, \infty) \subset E_n^c$. Let $f_n(x) = f(a) + \frac{x-a}{b-a} f(b)$ in the first case and $f_n(x) = f(a)$ in the other two cases.)
I claim that $f_n \to f$ almost everywhere. Suppose x is a point at which $f_n \not\to f$. Then $x \in (B_n^c) \cup (B_n \setminus E_n)$ for infinitely many n since otherwise $f_n(x)$ is eventually equal to $f(x)$. Now a given x can be in only finitely many B_n^c, so it must be in infinitely many $(B_n \setminus E_n)$, i.e. $x \in \limsup(B_n \setminus E_n)$. But $\limsup(B_n \setminus E_n)$ has measure 0 by the Borel-Cantelli Lemma. Hence the set of x at which $f_n(x) \not\to f(x)$ is a subset of a set of measure 0, and therefore has measure 0. □

Problem 20: Show that there exist closed sets A and B with $m(A) = m(B) = 0$, but $m(A + B) > 0$:

(a) In $\mathbb{R}$, let $A = \mathcal{C}$, $B = \mathcal{C}/2$. Note that $A + B \supset [0, 1]$.

(b) In $\mathbb{R}^2$, observe that if $A = I \times \{0\}$ and $B = \{0\} \times I$ (where $I = [0, 1]$), then $A + B = I \times I$.

Solution.

(a) As noted, let $\mathcal{C}$ be the Cantor set, $A = \mathcal{C}$, and $B = \mathcal{C}/2$. Then A consists of all numbers which have a ternary expansion using only 0's and 2's, as shown on a previous homework set. This implies that B consists of all numbers which have a ternary expansion using only 0's and 1's. Now any number $x \in 0, 1]$ can be written as $a + b$ where $a \in A$ and $b \in B$ as follows: Pick any ternary expansion $0.x_1x_2\dots$ for x. Define

$$a_n = \begin{cases} 2 & (x_n = 2) \\ 0 & (\text{else}) \end{cases}$$

$$b_n = \begin{cases} 1 & (x_n = 1) \\ 0 & (\text{else}) \end{cases}$$

Then $a = 0.a_1a_2\cdots \in A$ and $b = 0.b_1b_2\cdots \in B$, and $a+b = x$.

(b) Duly noted. (I feel like I should prove something, but I'm not sure what there is to prove here.)

□

Problem 21: Prove that there is a continuous function that maps a Lebesgue measurable set to a non-measurable set.

Proof. As shown in the last homework, there is a continuous function $f : [0,1] \to [0,1]$ such that $f(C) = [0,1]$, where C is the Cantor set. Let $V \subset [0,1]$ be a Vitali set, which is non-measurable. Let $E = f^{-1}(V) \cap C$. Then E is measurable since E is a subset of a set of measure 0. However, $f(E) = V$ which is not measurable. □

Problem 22: Let $\chi_{[0,1]}$ be the characteristic function of $[0,1]$. Show that there is no everywhere continuous function f on $\mathbb{R}$ such that

$$f(x) = \chi_{[0,1]}(x) \qquad \text{almost everywhere.}$$

Proof. Suppose that such an f exists. Then $f(1) = \chi_{[0,1]}(1) = 1$. By continuity, $\exists \delta > 0$ such that $|x-1| < \delta \Rightarrow |f(x) - 1| < \frac{1}{2}$. In particular, $f(x) > \frac{1}{2} > 0$ for $x \in (1, 1+\delta)$. Thus

$$\{x : f(x) \neq \chi_{[0,1]}(x)\} \supset (1, 1+\delta) \Rightarrow m(\{x : f(x) \neq \chi_{[0,1]}(x)) \geq \delta > 0.$$

□

Problem 23: Suppose $f(x,y)$ is a function on $\mathbb{R}^2$ that is separately continuous: for each fixed variable, f is continuous in the other variable. Prove that f is measurable on $\mathbb{R}^2$.

Solution. For all $x \in \mathbb{R}$ and $n \in \mathbb{N}$, we define D_x^n to be the largest nth-order dyadic rational less than or equal to x, i.e. $D_x^n = \frac{k}{2^n}$ where $\frac{k}{2^n} \leq x < \frac{k+1}{2^n}$. Let $f_n(x,y) = f(D_x^n, y)$. I will show that f_n is measurable and $f_n \to f$ everywhere.

First we show that f_n is measurable.

$$\begin{aligned}\{(x,y) : f_n(x,y) > a\} &= \bigcup_{k=-\infty}^{\infty} \left\{(x,y) : \frac{k}{2^n} \leq x < \frac{k+1}{2^n}, f\left(\frac{k}{2^n}, y\right) > a\right\} \\ &= \bigcup_{k=-\infty}^{\infty} \left[\frac{k}{2^n}, \frac{k+1}{2^n}\right) \times \left\{y : f\left(\frac{k}{2^n}, y\right) > a\right\}.\end{aligned}$$

Now because f is continuous in y, $\{y : f(\frac{k}{2^n}, y) > a\}$ is open, so the product $[\frac{k}{2^n}, \frac{k+1}{2^n}) \times \{y : f(\frac{k}{2^n}, y) > a\}$ is measurable. Hence $\{f_n > a\}$ is a countable union of measurable sets, and thus measurable.

Next, we show that $f_n \to f$ everywhere. Let $\epsilon > 0$. For any x, y, because f is continuous in x, there is some 1-dimensional neighborhood $(x \pm \delta, y)$ on which $|f(x', y) - f(x,y)| < \epsilon$. Then for sufficiently large n, $|D_x^n - x| < \delta \Rightarrow |f_n(x,y) - f(x,y)| = |f(D_x^n, y) - f(x,y)| < \epsilon$. Hence $f_n(x,y)$ is within ϵ of $f(x,y)$ for sufficiently large n. Since f is the pointwise limit of measurable functions f_n, it is measurable. □

Problem 27: Suppose E_1 and E_2 are a pair of compact sets in $\mathbb{R}^d$ with $E_1 \subset E_2$, and let $a = m(E_1)$ and $b = m(E_2)$. Prove that for any c with $a < c < b$, there is a compact set E with $E_1 \subset E \subset E_2$ and $m(E) = c$.

Solution. Since E_1 is measurable, there is an open set $U \supset E_1$ with $m(U \setminus E_1) < b - c$. Then $E_2 \cap U^c$ is compact (since it's the intersection of a compact set and a closed set) and has measure at least $m(E_1) - m(U) > b - (a + b - c) = c - a$. If we can find a compact subset $K \subset E_2 \cap U^c$ with $m(K) = c - a$, then $K \cup E_1$ will be a compact subset of E_2 with measure $(c - a) + a = c$ (since K and E_1 are disjoint). Hence we have reduced the problem to the following: Given a compact set $F \subset \mathbb{R}^d$ with $m(K) = \mu$, and given ξ with $0 < \xi < \mu$, find a compact subset $F' \subset F$ with $m(F') = \xi$. This can be solved as follows: Let $f(y) = m(F \cap B_y(0))$. Then $f(0) = 0$ whereas $f(y) = \mu$ for sufficiently large y (because F is bounded). Moreover, f is continuous: Given y and $\epsilon > 0$, the continuity of $m(B_y(0))$ allows us to find δ such that $|y' - y| < \delta \Rightarrow |m(B_{y'}(0)) - m(B_y(0))| < \epsilon$. Then $|f(y') - f(y)| = m(F \cap (B_{y'}(0)\Delta B_y(0))) < \epsilon$, because this is the measure of a subset of the symmetric difference $(B_{y'}(0)\Delta B_y(0))$ which has measure $|m(B_{y'}(0)) - m(B_y(0))| < \epsilon$. Hence f is continuous, so by the Intermediate Value Theorem there is a value y_0 such that $f(y_0) = \xi$. Then the compact set $F \cap B_{y_0}(0)$ has measure ξ as desired. □

Problem 28: Let $E \subset \mathbb{R}$ with $m_*(E) > 0$. Let $0 < \alpha < 1$. Then there exists an interval $I \subset \mathbb{R}$ such that $m_*(I \cap E) \geq \alpha m(I)$.

Proof. For any ϵ, we can find a cubical covering $\{Q_j\}$ of E with $\sum |Q_j| < m_*(E) + \epsilon$. Then, by expanding each cube to an open cube of size $\frac{\epsilon}{2^j}$ more, we can construct an *open* cubical covering $\{I_j\}$ with $\sum |I_j| < m_*(E) + 2\epsilon$. (We name these I_j because 1-dimensional cubes are in fact intervals.) Then

$$E \subset \bigcup_j I_j \Rightarrow E = \bigcup_j (E \cap I_j) \Rightarrow m_*(E) \leq \sum_j m_*(E \cap I_j).$$

Now we apply something like the Pigeonhole Principle: Suppose $m_*(E \cap I_j) < \alpha m(I_j)$ for all j. Then

$$m_*(E) \leq \sum_j m_*(E \cap I_j) < \alpha \sum_j m(I)_j < \alpha(m_*(E) + 2\epsilon).$$

But if ϵ is chosen small enough, this is a contradiction; explicitly, we can choose $\epsilon < \frac{(1-\alpha)m_*(E)}{2}$. Thus there must be some interval I_j for which $m_*(E \cap I_j) \geq \alpha m(I_j)$. □

Problem 29: Suppose E is a measurable subset of $\mathbb{R}$ with $m(E) > 0$. Prove that the difference set of E

$$\{z \in \mathbb{R} : z = x - y \text{ for some } x, y \in E\}$$

contains an open interval centered at the origin.

Solution. By Problem 28, there exists an interval I such that $m(E \cap I) \geq \frac{3}{4}m(I)$. (We can replace m_* by m here because E is measurable.) Let $d = m(I)$ and $0 < |\alpha| < \frac{d}{4}$. Then the translated set $I + \alpha$ intersects I in an

interval I' of length $d-\alpha > \frac{3}{4}d$. Now $m(E\cap I) \geq \frac{3}{4}d$ so $m((E+\alpha)\cap(I+\alpha)) \geq \frac{3}{4}d$ by the translation invariance of Lebesgue measure. Now

$$\frac{3}{4}d \leq m(E\cap I) = m(E\cap I')+m(E\cap(I\setminus I')) \leq m(E\cap I')+m(I\setminus I') = m(E\cap I')+\alpha$$

so $m(E\cap I') \geq \frac{3d}{4}-\alpha > \frac{d}{2}$. Similarly, $m((E+\alpha)\cap I') > \frac{d}{2}$. Now if $(E+\alpha)$ and E were disjoint, this would imply $m(I') \geq m(E\cap I')+m((E+\frac{d}{4})\cap I') > d$. But $m(I') = d-\alpha$, so E and $E+\alpha$ must have nonempty intersection. Let $x \in E\cap(E+\alpha)$. Then $\exists e_1, e_2 \in E$ such that $e_1 = x = e_2+\alpha \Rightarrow e_1-e_2 = \alpha$. Hence $\alpha \in E-E$ for all $\alpha \in (-\frac{d}{4}, \frac{d}{4})$. □

Problem 30: If E and F are measurable, and $m(E) > 0$, $m(F) > 0$, prove that

$$E+F = \{x+y : x \in E, y \in F\}$$

contains an interval.

Solution. We follow the preceding proof almost exactly. By the lemma, there exist intervals I_1 and I_2 such that $m(E\cap I_1) \geq \frac{3}{4}m(I_1)$ and $m(-F\cap I_2) \geq \frac{3}{4}m(I_2)$. WLOG assume $m(I_2) \leq m(I_1)$. Then $\exists t_0$ such that $I_2+t_0 \subset I_1$. Let $d = m(I_2)$. Then for $0 < |\alpha| < \frac{d}{4}$, $I_2+t_0+\alpha$ intersects I_1 in an interval I' of length at least $d-\alpha > \frac{3d}{4}$. By the same argument as in problem 29, this implies that I_1 and $I_2+t_0+\alpha$ must have nonempty intersection; let x be a point of this intersection. Then $\exists e \in E, f \in F$ such that $e = x = -f+t_0+\alpha \Rightarrow e+f = t_0+\alpha$. Hence $E+F$ contains the interval $(t_0-\frac{d}{4}, t_0+\frac{d}{4})$. □

Problem 34: Given two Cantor sets $\mathcal{C}_1, \mathcal{C}_2 \subset [0,1]$, there exists a continuous bijection $f : [0,1] \to [0,1]$ such that $f(\mathcal{C}_1) = \mathcal{C}_2$.

Proof. Any Cantor set can be put in bijective correspondence with the set of 0-1 sequences as follows: Given $x \in C$, where $C = C_1 \cap C_2 \cap \ldots$ is a Cantor set, define $x_1 = 0$ if x is in the left of the two intervals in C_1 (call this left interval I_0), and $x_1 = 1$ if x is in the right interval I_1. Then define $x_2 = 0$ if x is in the left subinterval (either I_{00} or I_{10}) in C_2, and $x_2 = 1$ if x is in the right subinterval. Continuing in this fashion, we obtain a bijection from C to the set of 0-1 sequences. Note that this bijection is increasing in the sense that if $y > x$ for $x, y \in C$, then $y_n > x_n$ at the first point n in the sequence at which x_n and y_n differ.
Now we can create an increasing bijection f from $\mathcal{C}_1$ to $\mathcal{C}_\in$ by mapping from $\mathcal{C}_1$ to 0-1 sequences, and from there to $\mathcal{C}_2$. This function will be continuous on $\mathcal{C}_1$ because if $x, y \in \mathcal{C}_1$ are close, their corresponding sequences $\{x_n\}$ and $\{y_n\}$ will agree in their first N terms; then $f(x)$ and $f(y)$ will agree in their first N terms as well, which means they're in the same subinterval of the Nth iterate of $\mathcal{C}_2$, which has length at most $\frac{1}{2^N}$. Hence $f(x)$ and $f(y)$ can be made arbitrarily close if x and y are sufficiently close. Then since $\mathcal{C}_1$ is compact, we can extend f to a continuous bijection on all of $[0,1]$ in a piecewise linear fashion, because $\mathcal{C}_1^c$ is a disjoint union of open intervals on which f can be made piecewise linear. This construction will also preserve the bijectivity of f. Hence we have a continuous bijection $f : [0,1] \to [0,1]$ with $f(\mathcal{C}_1) = \mathcal{C}_2$. □

Problem 35: Give an example of a measurable functions f and a continuous function Φ so that $f \circ \Phi$ is non-measurable. Use the construction in the hint to show that there exists a Lebesgue measurable set that is not a Borel set.

Solution. We will make use of Problem 34. Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be two Cantor subsets of $[0,1]$ such that $m(\mathcal{C}_1) > 0$ and $m(\mathcal{C}_2) = 0$. Let $N \subset \mathcal{C}_1$ such that N is non-measurable. (Here we use the fact that every set $E \subset \mathbb{R}$ of positive measure has a non-measurable subset. This is easy to prove by mimicking the Vitali construction but restricting it to E: Define the equivalence relation on E by x y if $x - y \in \mathbb{Q}$, and let $N \subset E$ contain one member from each equivalence class. Then N and its countably many translates are all of E, so N cannot have either measure 0 or nonzero measure.) Define $f = \chi_{\Phi(N)}$. Then $f \circ \Phi = \chi_N$ is non-measurable, since $\{x : \chi_N(x) > \frac{1}{2}\} = N$.
To show that there is a Lebesgue-measurable set which is not Borel measurable, we will use (without proof) the fact that every Borel set can be represented by a finite number of union and intersection signs, followed by some open sets. We also use the general fact that for any function ϕ and any sets X_α, $\phi^{-1}(\cap X_\alpha) = \cap \phi^{-1} X_\alpha$ and $\phi^{-1}(\cup X_\alpha) = \cup \phi^{-1} X_\alpha$. Together, these imply that for a continuous function Φ, Φ^{-1} of a Borel set is Borel, because

$$\Phi^{-1}\left(\bigcup \bigcap \cdots \bigcap G_n\right) = \bigcup \bigcap \cdots \bigcap \Phi^{-1}(G_n)$$

and $\Phi^{-1}(G_n)$ is open for continuous Φ and open G_n. (Of course, the string of cups and caps in this equation could just as well start with a cap.) Now consider the set $\Phi(N)$ from our construction above. Since $\Phi(N)$ is a subset of the set $\mathcal{C}_2$ which has measure 0, $\Phi(N)$ is measurable by the completeness of Lebesgue measure. However, Φ is a bijection so $\Phi^{-1}(\Phi(N)) = N$ which is not Borel, so $\Phi(N)$ cannot be Borel. □

1. Chapter 1.7, Page 46

Problem 2: Any open set U can be written as the union of closed cubes, so that $U = \cup Q_j$ with the following properties:

(i) The Q_j have disjoint interiors.

(ii) $d(Q_j, U^c) \approx$ side length of Q_j. This means that there are positive constants c and C so that $c \leq d(Q_j, U^c)/\ell(Q_j) \leq C$, where $\ell(Q_j)$ denotes the side length of Q_j.

Proof. Let $U \subset \mathbb{R}^d$ be an open set. Let $\mathcal{P}_0 = \{Q_j\}$ be the partition of U into dyadic cUbes as described in the book (pp. 7-8). I claim that $\frac{d(Q_j, U^c)}{\ell(Q_j)}$ is bounded above for this partition (i.e. does not take on arbitrarily large valUes). This is so because if $d(Q_j, U^c) > \sqrt{d}\ell(Q_j)$, then Q_j is a subset of a larger dyadic cube lying within U. (If we take the dyadic cube one size larger than Q_j and containing Q_j, then Q_j is at most $\sqrt{d}\ell_{Q_j}$ away from any point in this larger dyadic cube.) But the constrUction of $\mathcal{P}_0$ is done in such a way that every dyadic cube is maximal in U, i.e. is not contained in a larger dyadic cube lying within U. Hence $\frac{d(Q_j, U^c)}{\ell Q_j} \leq \sqrt{d}$ for all Q_j.

Now we define an iterative refining procedUre on the partition $\mathcal{P}_0$. In this procedure, every cube Q in the partition $\mathcal{P}_n$ for which $\frac{d(Q,U^c)}{\ell(Q)} < \frac{\sqrt{d}}{4}$ is replaced by its 2^d dyadic sub-cubes. If R is one of these sub-cubes, then

$$\frac{d(R,U^c)}{\ell(R)} \geq \frac{d(Q,U^c)}{\ell(Q)} = 2\frac{d(Q,U^c)}{\ell(Q)}$$

and

$$\frac{d(R,U^c)}{\ell(R)} \leq \frac{d + \sqrt{d}\ell(R)}{\ell(R)} = 2\frac{d(Q,U^c)}{\ell(Q)} + \sqrt{d}.$$

ThUs, if $\rho(Q) = \frac{d(Q,U^c)}{\ell(Q)}$, then the sub-cubes have the property

$$2\rho(Q) \leq \rho(R) \leq 2\rho(Q) + \sqrt{d}.$$

Now let

$$\mathcal{P} = \bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}\mathcal{P}_k$$

be the partition consisting of those cubes that are eventUally in all the $\mathcal{P}_n$. I claim that $\frac{\sqrt{d}}{4} \leq \frac{d(Q,U^c)}{\ell(Q)} \leq 2\sqrt{d}$ for any cube $Q \in \mathcal{P}$. Consider what happens as oUr refinement process iterates. If a given cube has too small a distance-to-side ratio, its sub-cubes will have this ratio at least doubled in the next iteration. Hence, after enough iterations its sub-cubes will all have their distance-to-side ratio in the desired interval $[\frac{\sqrt{d}}{4}, 2\sqrt{d}]$. Once they are in this interval, they are not sub-divided any further. One the other hand, none can ever achieve a ratio too large to be in this interval, since a cube is only subdivided if its ratio is less than $\sqrt{d}/4$, and the next iteration can then make it at most $2(\sqrt{d}/4) + \sqrt{d} = 3\sqrt{d}/2 < 2\sqrt{d}$. This shows that $\mathcal{P}$ has all its cubes in the desired interval. Consider also that $\mathcal{P}$ must have disjoint interiors and cover all of U: The only cubes with overlapping interiors are those from distinct steps in our iteration scheme, so taking the intersection to get $\mathcal{P}$ will weed out any such overlaps. Also, for a given $x \in U$, if we consider the sequence of cubes containing x in our various partitions, this sequence will shrink for a finite number of steps and then stay constant once the distance-to-side ratio reaches a desirable number. Hence x is contained in some cube that is eventually in all the $\mathcal{P}_n$, so x is covered by $\mathcal{P}$. □

Problem 3: Find an example of a measurable subset C of $[0,1]$ such that $m(C) = 0$, yet the difference set of C contains a non-trivial interval centered at the origin.

Solution. Let C be the Cantor middle-thirds set. Note that the Cantor dust described in the hint consists precisely of those point (x,y) for which both x and y are in C. Note also that if a line of slope 1 passes through any cube in any iteration of the Cantor dust, it must pass through one of the sub-cubes of that cube in the next iteration. To see this, consider WLOG the original cube. For a line $y = x + a$, if $\frac{1}{3} \leq a \leq 1$ then the line passes through the upper left cube; if $-\frac{1}{3} \leq a \leq \frac{1}{3}$ then it passes through the lower left cube; and if $-1 \leq a \leq -\frac{1}{3}$ it passes through the lower right cube. Thus, if C_n is the nth iteration of the Cantor dust and L is a line

of the form $y = x + a$ with $-1 \le a \le 1$, then $L \cap C_n$ is nonempty and compact. (It is nonempty by the preceding remark, and compact because it is the intersection of the compact sets C_n and $L \cap ([0,1] \times [0,1])$.) Thus, the infinite intersection of these nested compact sets is nonempty, i.e. L intersects the Cantor dust at some point (x, y). Then since $x, y \in C$, we have $x - y = a$. So a is in the difference set of the Cantor set for any $a \in [-1, 1]$. □

Problem 4: Complete the following outline to prove that a bounded function on an interval $[a, b]$ is Riemann integrable if and only if its set of discontinuities has measure zero. This argument is given in detail in the appendix to Book I.

Let f be a *bounded* function on a compact interval J, and let $I(c, r)$ denote the open interval centered at c of radius $r > 0$. Let $\text{osc}(f, c, r) = \sup |f(x) - f(y)|$ where the supremum is taken over all $x, y \in J \cap I(c, r)$, and define the oscillation of f at c by $\text{osc}(f, c) = \lim_{r \to 0} \text{osc}(f, c, r)$. Clearly, f is continuous at $c \in J$ if and only if $\text{osc}(f, c) = 0$.

Prove the following assertions:

(a) For every $\epsilon > 0$, the set of points $c \in J$ such that $\text{osc}(f, c) \ge \epsilon$ is compact.

(b) If the set of discontinuities of f has measure 0, then f is Riemann integrable.

(c) Conversely, if f is Riemann integrable on J, then its set of discontinuities has measure 0.

Solution.

(a) Clearly this set is bounded, so we need only show that it's closed. Suppose $c_1, c_2, \ldots$ are a sequence of such points with $c_n \to c$. We wish to show that $\text{osc}(f, c) \ge \epsilon$ as well. Let $r > 0$. Since $c_n \to c$, there is some c_N with $|c_N - c| < \frac{r}{2}$. Then since $\text{osc}(f, c_N) \ge \epsilon$, there must be $x, y \in J$ within $\frac{r}{2}$ of c_N such that $|f(x) - f(y)| \ge \epsilon$. But then x and y are within r of c, so $\text{osc}(f, c, r) \ge \epsilon$. Since this is true for all r, we must have $\text{osc}(f, c) \ge \epsilon$.

(b) Let $\epsilon > 0$. Let

$$A_\epsilon = \{c \in J : \text{osc}(f, c) \ge \epsilon\}.$$

Since A_ϵ is a subset of the discontinuity set of f, it has measure 0. Hence there it can be covered by an open set U with $m(U) < \epsilon$. Since every open subset of $\mathbb{R}$ is a countable disjoint union of intervals, we can write $U = \cup I_n$ with $\sum |I_n| < \epsilon$. Now because A_ϵ is compact, there is a finite subcover; by re-ordering the intervals we may write

$$A_\epsilon \subset \bigcup_{n=1}^{N} I_n.$$

Now on the compact set $J' = J \setminus \cup I_n$, $\text{osc}(f, c) < \epsilon$ for all c. For each $x \in J'$, this means we can find r_x such that $\text{osc}(f, x, r_x) < \epsilon$. Then J' is covered by the open intervals $U_x = (x - r_x, x + r_x)$. Let δ be the Lebesgue number of this covering, so that any subinterval of J' with length at most δ must be contained in one of the U_x. Now

consider a partition of J with mesh size less than δ. The total length of all subintervals which intersect $\cup I_n$ is at most $\epsilon + 2N\delta$ since enlarging each I_n by δ will cover all such intervals. On each of these subintervals, $\sup f - \inf f \leq 2M$ where $|f| \leq M$ on J. Hence the contribution these intervals make to the difference $U(P, f) - L(P, f)$ is at most $2M\epsilon + 4M\delta$. The other subintervals are contained in J' and by construction of δ, each is contained within some U_x, so $\sup f - \inf f \leq \epsilon$ on each of them. Hence the total contribution they make to $U(P, f) - L(P, f)$ is at most $\epsilon m(J)$. Thus, we have

$$U(P, f) - L(P, f) \leq \epsilon(2M + m(J) + 4\delta).$$

By requiring δ to be less than some constant times ϵ, we have thus shown that the difference between upper and lower sums can be made smaller than a constant times ϵ. Hence f is Riemann integrable.

(c) Suppose f is Riemann integrable, and let $\epsilon > 0$. Let $n \in \mathbb{N}$. Then there is a partition P of J with $U(P, f) - L(P, f) < \frac{\epsilon}{n}$. Now if the interior of any subinterval I_k of this partition intersects $A_{1/n}$ at some x, then $\sup f - \inf f \geq \frac{1}{n}$ on I_k because $\text{osc}(f, x, r) \geq \frac{1}{n}$ for all r, and $(x - r, x + r) \subset I_k$ for sufficiently small r. So the total length of the subintervals whose interiors intersect $A_{1/n}$ is at most ϵ since otherwise they would make a contribution of more than ϵ/n to $U(P, f) - (P, f)$. Hence we have covered $A_{1/n}$ by a collection of intervals of total length less than ϵ, which implies $m(A_{1/n}) < \epsilon$. Now if A is the set of points at which f is discontinuous, then

$$A = \bigcup_{n=1}^{\infty} A_{1/n}.$$

Since $A_n \subset A_{n+1}$, the continuity of measure implies that

$$m(A) = \lim_{n \to \infty} m(A_{1/n}) \leq \epsilon.$$

Since $m(A) \leq \epsilon$ for all ϵ, $m(A) = 0$.

□

Problem 6: The fact that the axiom of choice and the well-ordering principle are equivalent is a consequence of the following considerations.

One begins by defining a partial ordering on a set E to be a binary relation $\leq$ on the set E that satisfies:

(i) $x \leq x$ for all $x \in E$.
(ii) If $x \leq y$ and $y \leq x$, then $x = y$.
(iii) If $x \leq y$ and $y \leq z$, then $x \leq z$.

If in addition $x \leq y$ or $y \leq x$ whenever $x, y \in E$, then $\leq$ is a linear ordering of E.

The axiom of choice and the well-ordering principle are then logicall equivalent to the Hausdorff maximal principle: Every non-empty partially ordered set has a (non-empty) maximal linearly ordered subset. In other words, if E is partially ordered by $\leq$, then E contains a non-empty subset F which is linearly ordered by $\leq$ and such that if F is contained in a set G also linearly ordered by $\leq$, then $F = G$.

An application of the Hausdorff maximal principle to the collection of all

well-orderings of subsets of E implies the well-ordering principle for E. However, the proof that the axiom of choice implies the Hausdorff maximal principle is more complicated.

Solution. I don't like this problem for two reasons: (1) It's not very clearly stated what exactly I'm supposed to do. My best guess is that I'm supposed to use the axiom of choice to prove the Hausdorff maximal principle, but the problem never really comes out and says that. (2) What the crap. This is supposed to be an analysis course, not a set theory course. We haven't defined any of the basic concepts of set theory, yet I'm supposed to come up with a "complicated" set theory proof. Since I've never had a course in set theory, I really don't feel like guessing my way to something that might be a proof. So, here's the proof from the appendix to Rudin's *Real and Functional Analysis*:
For a collection of set $\mathcal{F}$ and a sub-collection $\Phi \subset \mathcal{F}$, call Φ a *subchain* of $\mathcal{F}$ if Φ is totally ordered by set inclusion.

Lemma 1. *Suppose $\mathcal{F}$ is a nonempty collection of subsets of a set X such that the union of every subchain of $\mathcal{F}$ belongs to $\mathcal{F}$. Suppose g is a function which associates to each $A \in \mathcal{F}$ a set $g(A) \in \mathcal{F}$ such that $A \subset g(A)$ and $g(A) \setminus A$ consists of at most one element. Then there exists an $A \in \mathcal{F}$ for which $g(A) = A$.*

Proof. Let $A_0 \in \mathcal{F}$. Call a subcollection $\mathcal{F}' \subset \mathcal{F}$ a *tower* if $A_0 \in \mathcal{F}'$, the union of every subchain of $\mathcal{F}'$ is in $\mathcal{F}'$, and $g(A) \in \mathcal{F}'$ for every $A \in \mathcal{F}'$. Then there exists at least one tower because the collection of all $A \in \mathcal{F}$ such that $A_0 \subset A$ is a tower. Let $\mathcal{F}_0$ be the intersection of all towers, which is also a tower. Let Γ be the collection of all $C \in \mathcal{F}_0$ such that every $A \in \mathcal{F}_0$ satisfies either $A \subset C$ or $C \subset A$. For each $C \in \Gamma$, let $\Phi(C)$ be the collection of $A \in \mathcal{F}_0$ such that $A \subset C$ or $g(C) \subset A$. We will prove that Γ is a tower. The first two properties are obvious. Now let $C \in \Gamma$, and suppose $A \in \Phi(C)$. If A is a proper subset of C, then C cannot be a proper subset of $g(A)$ because then $g(A) \setminus A$ would have at least two elements. Hence $g(A) \subset C$. If $A = C$, then $g(A) = g(C)$. The third possibility for A is that $g(C) \subset A$. But since $A \subset g(A)$ this implies $g(C) \subset g(A)$. Thus, we have shown that $g(A) \in \Phi(C)$ for any $A \in \Phi(C)$, so $\Phi(C)$ is a tower. By the minimality of $\mathcal{F}_0$, we must have $\Phi(C) = \mathcal{F}_0$ for every $C \in \Gamma$. This means that $g(C) \in \Gamma$ for all $C \in \Gamma$, so Γ is also a tower; by minimality again, $\Gamma = \mathcal{F}_0$. This shows that $\mathcal{F}_0$ is totally ordered.
Now let A be the union of all sets in $\mathcal{F}_0$. Then $A \in \mathcal{F}_0$ by the second tower property, and $g(A) \in \mathcal{F}_0$ by the third. But A is the largest member of $\mathcal{F}_0$ and $A \subset g(A)$, so $A = g(A)$. $\square$

Now let $\mathcal{F}$ be the collection of all totally ordered subsets of a partially ordered set E. Since every single-element subset of E is totally ordered, $\mathcal{F}$ is not empty. Note that the union of any chain of totally ordered sets is totally ordered. Now let f be a choice function for E. If $A \in \mathcal{F}$, let A^* be the set of all x in the complement of A such that $A \cup \{x\} \in \mathcal{F}$. If $A^* \neq \emptyset$, let

$$g(A) = A \cup \{f(A^*)\}.$$

If $A^* = \emptyset$, let $g(A) = A$.
By the lemma, $A^* = \emptyset$ for at least one $A \in \mathcal{F}$, and any such A is a maximal element of $\mathcal{F}$. $\square$

Problem 7: Consider the curve $\Gamma = \{y = f(x)\}$ in $\mathbb{R}^2$, $0 \leq x \leq 1$. Assume that f is twice continuously differentiable in $0 \leq x \leq 1$. Then show that $m(\Gamma + \Gamma) > 0$ if and only if $\Gamma + \Gamma$ contains an open set, if and only if f is not linear.

Solution. We are asked to show the equivalence of the conditions (i) $m(\Gamma + \Gamma) > 0$, (ii) $\Gamma + \Gamma$ contains an open set, and (iii) f is not linear. We will show that (ii) implies (i), which implies (iii), which implies (ii).
First, we should note that $\Gamma + \Gamma$ is measurable. The problem doesn't ask for this, but it's worth pointing out. Consider $G : [0, 1] \times [0, 1] \to \mathbb{R}^2$ defined by $G(x, y) = (x + y, f(x) + f(y))$. Then $\Gamma + \Gamma$ is just the range of G. Since differentiable functions map measurable sets to measurable sets, $\Gamma + \Gamma$ is measurable. (We haven't proved yet that differentiable functions preserve measurability, but I assume we will once we get further into differentiation theory.)
The easiest of our three implications is (ii) implies (i). Suppose $\Gamma + \Gamma$ contains an open set. Open sets have positive measure, so $\Gamma + \Gamma$ is a measurable set with a subset of positive measure, so it has positive measure.
Now suppose $m(\Gamma + \Gamma) > 0$. We wish to show that f is not linear. Suppose instead that f is linear, say $f(x) = ax + b$. Then for any x, x', $(x + x', f(x) + f(x')) = (x + x', a(x + x') + 2b)$ so $\Gamma + \Gamma$ is a subset of the line $y = ax + 2b$, which has measure 0. Thus, if $\Gamma + \Gamma$ has positive measure, f must not be linear.
The third implication is the least trivial. Suppose f is not linear. Then there are points $x_0, y_0 \in [0, 1]$ with $f'(x_0) \neq f'(y_0)$. Then the Jacobian

$$DG = \begin{vmatrix} 1 & 1 \\ f'(x) & f'(y) \end{vmatrix} = f'(y) - f'(x)$$

is nonzero at the point $(x, y) \in [0, 1] \times [0, 1]$, where $G(x, y) = (x + y, f(x) + f(y))$ as above. WLOG we may assume $(x, y) \in (0, 1) \times (0, 1)$ since a nonlinear function on $[0, 1]$ with continuous derivative cannot have constant derivative everywhere on $(0, 1)$. Then the Inverse Function Theorem guarantees that there is an open neighborhood of (x, y) on which G is a diffeomorphism; since diffeomorphisms are homeomorphisms, this implies that the image of G contains an open set. $\square$

Chapter 2.5, Page 89

Exercise 1: Given a collection of sets $F_1, \ldots, F_n$, construct another collection $F_1^*, \ldots, F_N^*$, with $N = 2^n - 1$, so that $\bigcup_{k=1}^n F_k = \bigcup_{j=1}^N F_j^*$; the collection $\{F_j^*\}$ is disjoint; and $F_k = \bigcup_{F_j^* \subset F_k} F_j^*$ for every k.

Solution. For $j = 1, \ldots, N$, create F_j^* as follows: First write j as an n-digit binary number $j_1 j_2 \ldots j_n$. Then for $k = 1, \ldots, n$, let

$$G_k = \begin{cases} F_k & j_k = 1 \\ F_k^c & j_k = 0. \end{cases}$$

Finally, let

$$F_j^* = \bigcap_{k=1}^n G_k.$$

For example, $2 = 000\dots 10$ in binary, so

$$F_2^* = F_1^c \cap F_2^c \cap \dots \cap F_{n-2}^c \cap F_{n-1} \cap F_n^c.$$

Note that the F_j^* are pairwise disjoint because if $j \neq j'$, then they differ in some binary digit, say $j_\ell \neq j'_\ell$. Suppose WLOG that $j_\ell = 1$ and $j'_\ell = 0$. Then $F_j^* \subset F_\ell$ whereas $F_{j'}^* \subset F_\ell^c$, so they are disjoint.
Also,

$$F_k = \bigcup_{F_j^* \subset F_k} F_j^*.$$

To see this, note that the RHS is clearly a subset of the LHS since it is a union of subsets. Conversely, suppose $x \in F_k$. Define $x_1, \dots, x_n$ by $x_i = 1$ if $x \in F_i$ and 0 otherwise. Then if m has the binary digits $m_1 = x_1, \dots, m_n = x_n$, $x \in F_m^*$ by definition of F_m^*. Since $F_m^* \subset F_k$, the result follows.
This implies

$$\bigcup_{i=1}^n F_i \subset \bigcup_{j=1}^N F_j^*.$$

But $F_j^* \subset \bigcup F_i$ for each j, so the reverse inclusion holds as well. $\square$

Exercise 2: In analogy to Proposition 2.5, prove that if f is integrable on $\mathbb{R}^d$ and $\delta > 0$, then $f(\delta x)$ converges to $f(x)$ in the L^1 norm as $\delta \to 1$.

Solution. Let $\epsilon > 0$. Since $C_C(\mathbb{R}^d)$ is dense in $L^1(\mathbb{R}^d)$, we can choose $g \in C_C(\mathbb{R}^d)$ such that $\|f - g\|_1 < \frac{\epsilon}{3}$. We can also choose λ such that $|\delta - 1| < \lambda \Rightarrow \frac{1}{\delta^d} < \frac{3}{2}$. Now let $K = \text{supp}(g)$, and choose M such that $K \subset \{|x| \leq M\}$. Let $B = \{|x| \leq M + 1\}$, which is also compact. Because g is uniformly continuous on B, $\exists \lambda' > 0$ such that $\lambda' < 1$ and for $x, y \in B$, $|x - y| < \lambda' \Rightarrow |g(x) - g(y)| < \frac{\epsilon}{6m(B)}$. Now suppose we choose δ such that $|\delta - 1| < min(\lambda, \frac{\lambda'}{M+1})$. Then

$$\begin{aligned} \|f(\delta x) - f(x)\| &= \int |f(\delta x) - f(x)| \\ &\leq \int |f(\delta x) - g(\delta x)| + \int |f(\delta x) - g(x)| + \int |g(x) - f(x)| \end{aligned}$$

by the triangle inequality. The third integral is just $\|f - g\| \leq \frac{\epsilon}{3}$. By the dilation property of the integral, the first integral is $\frac{1}{\delta^d}\|f - g\| \leq \frac{3}{2}\frac{\epsilon}{3} = \frac{\epsilon}{2}$. Now for the second integral. I claim that the integrand is 0 outside B and at most $\frac{\epsilon}{6m(B)}$ inside. If $x \notin B$, then $|x| > M + 1$ so

$$|\delta x| = |\delta||x| \geq \left(1 - \frac{\lambda'}{M+1}\right)|x| > \left(1 - \frac{\lambda'}{M+1}\right)(M+1) > M,$$

which implies that $g(\delta x) = g(x) = 0$. Now suppose $x \in B$. If $g(x) \neq 0$, then $x \in K$ and

$$|\delta x| \leq \left(1 + \frac{\lambda'}{M+1}\right)|x| < \left(1 + \frac{\lambda'}{M+1}\right)M < M + 1$$

so $\delta x \in B$ and $|\delta x - x| < \lambda' \Rightarrow |g(\delta x) - g(x)| < \frac{\epsilon}{6m(B)}$. If $g(\delta x) \neq 0$, then $\delta x \in K \subset B$ and again $|g(\delta x) - g(x)| < \frac{\epsilon}{6m(B)}$ because of uniform continuity. Hence the integrand is 0 outside B and is at most $\frac{\epsilon}{6m(B)}$ inside B, so its integral is at most $\frac{\epsilon}{6m(B)} m(B) = \frac{\epsilon}{6}$. Putting the pieces together, we have $\|f(\delta x) - f(x)\| < \epsilon$. □

Exercise 4: Suppose f is integrable on $[0, b]$, and

$$g(x) = \int_x^b \frac{f(t)}{t} dt \quad \text{for } 0 < x \leq b.$$

Prove that g is integrable on $[0, b]$ and

$$\int_0^b g(x) dx = \int_0^b f(t) dt.$$

Solution. We may assume WLOG that $f(t) \geq 0$, since otherwise we can analyze f^+ and f^- separately. Now let

$$h(x,t) = \frac{f(t)}{t} \chi_{\{0 < x \leq t \leq b\}}.$$

Then $h \geq 0$ and h is clearly measurable since it is a quotient of measurable functions times another measurable function. By Tonelli's theorem,

$$\int_{-\infty}^{\infty} h(x,t) dt = \int_x^b \frac{f(t)}{t} \chi_{\{0 < x \leq b\}} dt$$

is a measurable function of x. Note that this is just equal to $g(x)$ for $0 < x \leq b$ and 0 elsewhere. Hence g is measurable (in general, $g : (0, b] \to \mathbb{R}$ is measurable iff $g \cdot \chi_{(0,b]} : \mathbb{R} \to \mathbb{R}$ is). Moreover, Tonelli's theorem also tells us that

$$\begin{aligned}
\int_0^b g(x)\, dx &= \int_0^b \left(\int_x^t \frac{f(t)}{t}\, dt \right) dx \\
&= \int_{\mathbb{R} \times \mathbb{R}} h(x,t) \\
&= \int_0^b \left(\int_0^t h(x,t)\, dx \right) dt \\
&= \int_0^b \left(\int_0^t \frac{f(t)}{t}\, dx \right) dt \\
&= \int_0^b t \frac{f(t)}{t}\, dt \\
&= \int_0^b f(t)\, dt.
\end{aligned}$$

Note that the fact that this integral is finite implies that g is integrable and not just measurable. □

Exercise 5: Suppose F is a closed set in $\mathbb{R}$, whose complement has finite measure, and let $\delta(x)$ denote the distance from x to F, that is,

$$\delta(x) = d(x, F) = \inf\{|x - y| : y \in F\}.$$

Consider

$$I(x) = \int_{\mathbb{R}} \frac{\delta(y)}{|x-y|^2} dy.$$

(a) Prove that δ is continuous by showing that it satisfies the Lipschitz condition

$$|\delta(x) - \delta(y)| \leq |x - y|.$$

(b) Show that $I(x) = \infty$ for each $x \notin F$.
(c) Show that $I(x) < \infty$ for a.e. $x \in F$. This may be surprising in view of the fact that the Lipschitz condition cancels only one power of $|x - y|$ in the integrand of I.

Solution.
(a) Let $\epsilon > 0$. Choose $z \in F$ such that $|y - z| < \delta(y) + \epsilon$. Then

$$\delta(x) \leq |x-z| \leq |x-y| + |y-z| < |x-y| + \delta(y) + \epsilon \Rightarrow \delta(x) - \delta(y) < |x-y| + \epsilon.$$

Interchanging the roles of x and y, we also have

$$\delta(y) - \delta(x) < |x - y| + \epsilon.$$

Hence $|\delta(x) - \delta(y)| < |x - y| + \epsilon$ for any $\epsilon > 0$. This implies $|\delta(x) - \delta(y)| \leq |x - y|$.

(b) Suppose $x \notin F$. Because F is closed, this implies $\delta(x) > 0$, since otherwise there would be a sequence of points in F converging to x. Let $\lambda = \delta(x)$. By the Lipschitz condition from part (a), $|x - y| < \frac{\lambda}{2} \Rightarrow |\delta(y) - \lambda| < \frac{\lambda}{2} \Rightarrow \delta(y) \geq \frac{\lambda}{2}$. Hence

$$\begin{aligned} I(x) &= \int_{-\infty}^{\infty} \frac{\delta(y)}{|x-y|^2} \, dy \\ &\geq \int_{x-\lambda/2}^{x+\lambda/2} \frac{\delta(y)}{|x-y|^2} \, dy \\ &\geq \int_{x-\lambda/2}^{x+\lambda/2} \frac{\frac{\lambda}{2}}{|x-y|^2} \, dy \\ &= \frac{\lambda}{2} \int_{-\lambda/2}^{\lambda/2} \frac{1}{y^2} \, dy = \infty. \end{aligned}$$

(c) First, consider that for $y \notin F$,

$$\int_F \frac{1}{|x-y|^2} \, dx \leq 2 \int_{\delta(y)}^{\infty} \frac{1}{x^2} \, dx = \frac{2}{\delta(y)}$$

since $F \subset \{x : |x-y| \geq \delta(y)\}$. Now since $I(x) \geq 0$, we have by Tonelli's theorem

$$\begin{aligned}
\int_F I(x)\,dx &= \int_{\mathbb{R}\times\mathbb{R}} \frac{\delta(y)}{|x-y|^2}\chi_F(x) \\
&= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \frac{\delta(y)}{|x-y|^2}\chi_F(x)\,dx \right) dy \\
&= \int_{F^c} \delta(y) \left(\int_F \frac{1}{|x-y|^2}\,dx \right) dy \\
&\leq \int_{F^c} \delta(y) \frac{2}{\delta(y)}\,dy \\
&= 2m(F^c) < \infty.
\end{aligned}$$

Since $\int_F I(x)\,dx < \infty$, we must have $I(x) < \infty$ for almost all $x \in F$. (This is actually not all that shocking, since $I(x)$ is clearly less than ∞ for an interior point. Of course, there are closed sets whose boundaries have positive measure, but those are the nasty guys.)

□

Exercise 6: Integrability of f on $\mathbb{R}$ does not necessarily imply the convergence of $f(x)$ to 0 as $x \to \infty$.

(a) There exists a positive continuous function f on $\mathbb{R}$ such that f is integrable on $\mathbb{R}$, yet $\limsup_{x\to\infty} f(x) = \infty$.

(b) However, if we assume that f is uniformly continuous on $\mathbb{R}$ and integrable, then $\lim_{|x|\to\infty} f(x) = 0$.

Solution.

(a) Let

$$f(x) = \begin{cases} 2^{3n+4} d\left(x, \left[n, n + \frac{1}{2^{2n+1}}\right]^c\right) & \left(n \leq x \leq n + \frac{1}{2^{2n+1}}, n \in \mathbb{Z}\right) \\ 0 & \text{else.} \end{cases}$$

The graph of f consists of a series of triangular spikes with height 2^{n+2} and base $\frac{1}{2^{2n+1}}$. The nth such spike has area 2^{-n}, so $\int |f| = \sum_{n=0}^{\infty} 2^{-n} = 2$. But, because the spikes get arbitrarily high, $\limsup f(x) = \infty$.

(b) Suppose f is uniformly continuous on $\mathbb{R}$, and let $\epsilon > 0$. Select $\delta > 0$ such that $|x-y| < \delta \Rightarrow |f(x) - f(y)| < \frac{\epsilon}{2}$, and also require $\delta < \frac{1}{2}$. Since $f(x) \not\to 0$, $\exists x_1 > 0$ such that $|f(x_1)| \geq \epsilon$. Then $|f(y)| \geq \frac{\epsilon}{2}$ for $y \in (x_1 - \delta, x_1 + \delta)$. Now since $f(x) \not\to 0$, $\exists x_2 > x_1 + 1$ with $|f(x_2)| \geq \epsilon$. Then $|f| \geq \frac{\epsilon}{2}$ on $(x_2 - \delta, x_2 + \delta)$. Continuing in this manner, we obtain infinitely many intervals of length 2δ on which $|f| \geq \frac{\epsilon}{2}$. These intervals are disjoint because of our requirements that $|x_{n+1} - x| > 1$ and $\delta < \frac{1}{2}$. Hence, by Tchebycheff's inequality,

$$\int_{\mathbb{R}} |f(x)|\,dx \geq \frac{\epsilon}{2} m(\{x : |f(x)| \geq \frac{\epsilon}{2}\}) = \infty.$$

□

Exercise 8: If f is integrable on $\mathbb{R}$, show that

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

is uniformly continuous.

Solution. Let $\epsilon > 0$. By the absolute continuity of the integral (Prop 1.12b), $\exists \delta > 0$ such that $m(E) < \delta \Rightarrow \int_E |f| < \epsilon$. Then (assuming WLOG $x > y$),

$$|x - y| < \delta \Rightarrow |F(x) - F(y)| = \left| \int_y^x f(t)dt \right| \leq \int_y^x |f(t)|\, dt < \epsilon$$

because $m([y, x]) = |x - y| < \delta$. □

Exercise 10: Suppose $f \geq 0$, and let $E_{2^k} = \{x : f(x) > 2^k\}$ and $F_k = \{x : 2^k < f(x) \leq 2^{k+1}\}$. If f is finite almost everywhere, then

$$\bigcup_{k=-\infty}^{\infty} F_k = \{f(x) > 0\}$$

and the sets F_k are disjoint.
Prove that f is integrable if and only if

$$\sum_{k=-\infty}^{\infty} 2^k m(F_k) < \infty, \quad \text{if and only if } \sum_{k=-\infty}^{\infty} 2^k m(E_{2^k}) < \infty.$$

Use this result to verify the following assertions. Let

$$f(x) = \begin{cases} |x|^{-a} & \text{if } |x| \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

and

$$g(x) = \begin{cases} |x|^{-b} & \text{if } |x| > 1 \\ 0 & \text{otherwise.} \end{cases}$$

Then f is integrable on $\mathbb{R}^d$ if and only if $a < d$; also g is integrable on $\mathbb{R}^d$ if and only if $b > d$.

Solution. Let

$$g(x) = \sum_{k=-\infty}^{\infty} 2^k \chi_{F_k}(x),$$

$$h(x) = \sum_{k=-\infty}^{\infty} 2^{k+1} \chi_{F_k}(x).$$

Then $g(x) \leq f(x) \leq h(x)$ by definition of F_k. Then

$$\int f(x)dx < \infty \Rightarrow \int g(x)dx = \sum_{k=-\infty}^{\infty} 2^k m(F_k) < \infty$$

whereas

$$\sum_{k=-\infty}^{\infty} 2^k m(F_k) < \infty \Rightarrow \int f(x)dx < \int h(x)dx = \sum_{k=-\infty}^{\infty} 2^{k+1} m(F_k) = 2 \sum_{k=-\infty}^{\infty} 2^k m(F_k) < \infty.$$

Now let

$$\phi(x) = \sum_{k=-\infty}^{\infty} 2^k \chi_{E_k}(x).$$

Then $f(x) \le \phi(x) \le 2f(x)$ because if $2^k < f(x) \le 2^{k+1}$, $\phi(x) = \sum_{j=-\infty}^{k} 2^k = 1 + 1 + 2 + 4 + \cdots + 2^k = 2^{k+1}$. Hence

$$\int f(x)dx < \infty \Leftrightarrow \int \phi(x)dx = \sum_{k=-\infty}^{\infty} 2^k m(E_k) < \infty.$$

Now for the function f given,

$$E_k = \{f(x) > 2^k\} = \begin{cases} \{|x| \le 1\} & k \le 0 \\ \{|x| \le 2^{-k/a}\} & k \ge 1 \end{cases}$$

so

$$m(E_k) = \begin{cases} 2^d & k \le 0 \\ 2^d 2^{-kd/a} & k \ge 1 \end{cases}.$$

So f is integrable iff

$$\sum_{k=-\infty}^{\infty} 2^k m(E_k) = \sum_{k=-\infty}^{0} 2^k 2^d + \sum_{k=1}^{\infty} 2^k 2^d 2^{-kd/a} = 2^{d+1} + 2^d \sum_{k=-\infty}^{\infty} 2^{(1-d/a)k} < \infty.$$

This infinite sum will converge iff the constant $1 - \frac{d}{a}$ is negative, i.e. iff $a < d$.

For the function g given, let us redefine $g(x) = 1$ for $|x| \le 1$; clearly this does not affect the integrablity of g. Now E_k is empty for $k > 0$, so we need only consider negative values of k.

$$g(x) > 2^k \Leftrightarrow |x| < 2^{-k/b}$$

so E_k is a cube of volume $2^d 2^{-kd/b}$. Hence g is integrable iff

$$\sum_{k=-\infty}^{0} 2^k 2^d 2^{-kd/b} = 2^d \sum_{k=-\infty}^{0} 2^{(1-d/b)k}$$

converges. This will happen iff $1 - d/b > 0 \Rightarrow b > d$. □

Exercise 11: Prove that if f is integrable on $\mathbb{R}^d$, and $\int_E f(x)dx \ge 0$ for every measurable E, then $f(x) \ge 0$ a.e. x. As a result, if $\int_E f(x)dx = 0$ for every measurable E, then $f(x) = 0$ a.e.

Proof. Suppose it is not true that $f(x) \ge 0$ a.e., so $m(\{f(x) < 0\}$ is positive. Now

$$\{x : f(x) < 0\} = \bigcup_{n=1}^{\infty} \left\{x : f(x) < -\frac{1}{n}\right\} \Rightarrow m(\{x : f(x) < 0\}) \le \sum_{n=1}^{\infty} m\left(\left\{x : f(x) < -\frac{1}{n}\right\}\right)$$

by countable additivity. Hence at least one of the sets

$$E_n = \left\{x : f(x) < -\frac{1}{n}\right\}$$

has positive measure. But then

$$\int_{E_n} f(x)dx \le \int_{E_n} -\frac{1}{n}dx = -\frac{1}{n}m(E_n) < 0.$$

By contraposition, if $\int_E f(x)dx \geq 0$ for every measurable set E, then $f(x) \geq 0$ a.e.
Now if $\int_E f(x)dx = 0$ for every measurable E, then $\int_E f(x)dx \geq 0$ and $\int_E -f(x)dx \geq 0$, which means $f \geq 0$ a.e. and $-f \geq 0$ a.e. Hence $f = 0$ a.e. □

Exercise 12: Show that there are $f \in L^1(\mathbb{R}^d)$ and a sequence $\{f_n\}$ with $f_n \in L^1(\mathbb{R}^d)$ such that

$$\|f - f_n\|_1 \to 0,$$

but $f_n(x) \to f(x)$ for no x.

Solution. To assist in constructing such a sequence, we first construct a sequence of measurable sets $E_n \subset \mathbb{R}^d$ with the property that $m(E_n) \to 0$ but every $x \in \mathbb{R}^d$ is in infinitely many E_n. We proceed as follows: Choose integers $N_1, N_2, \ldots$ such that

$$1 + \frac{1}{2} + \cdots + \frac{1}{N_1} > 10$$
$$\frac{1}{N_1+1} + \frac{1}{N_1+2} + \cdots + \frac{1}{N_2} > 100$$
$$\frac{1}{N_2+1} + \frac{1}{N_2+2} + \cdots + \frac{1}{N_3} > 1000$$

etc. This is possible because of the divergence of the harmonic series. For convience, we also define $N_0 = 0$. Next, for each $k = 0, 1, 2, \ldots,$ let B_{N_k+1} be the cube of volume $\frac{1}{N_k+1}$ centered at the origin. Then, for a given k, define B_j for $N_k + 1 < j \leq N_{k+1}$ to be the cube centered at the origin with $|B_j| = |B_{j-1}| + \frac{1}{j}$. Finally, we define $E_{N_k+1} = B_{N_k+1}$, and for $N_k + 1 < j \leq N_{k+1}$, $E_j = B_j \setminus B_{j-1}$. I claim that the sets E_n have the desired properties. First, note that $m(E_n) = \frac{1}{n}$. This is obvious for $N_k + 1$; for $N_k+1 < j \leq N_{k+1}$ it is easy to see inductively that $|B_j| = \frac{1}{N_k+1} + \cdots + \frac{1}{j}$ and since they are nested sets, $|B_j \setminus B_{j-1}| \frac{1}{j}$. Thus $m(E_n) \to 0$. However, for each k,

$$\bigcup_{j=N_k+1}^{N_{k+1}} E_j$$

is a cube centered at the origin with a volume greater than 10^k. For any given x, these cubes will eventually contain x, i.e. there is some K such that

$$k > K \Rightarrow x \in \bigcup_{j=N_k+1}^{N_{k+1}} E_k.$$

Hence every x is in infinitely many E_j as desired.
Having constructed these sets, we simply let $f_n(x) = \chi_{E_n}(x)$ and $f(x) = 0$. Then $\int f_n(x)dx = \frac{1}{n}$ so $\|f_n - f\| = \int |f_n - f| = \int f_n \to 0$, i.e. $f_n \stackrel{L^1}{\to} f$. However, for any given x there are infinitely many n such that $f_n(x) = 1$, so $f_n(x) \not\to f(x)$ for any x. □

Exercise 13: Give an example of two measurable sets A and B such that $A + B$ is not measurable.

Solution. As suggested in the hint, let $N \subset \mathbb{R}$ be a non-measurable set. Let $A = \{0\} \times [0,1]$ and $B = N \times \{0\}$. Then A and B are both measurable subsets of $\mathbb{R}^2$ because they are subsets of lines, which have measure 0. Now $A+B = N \times [0,1]$. By Proposition 3.4, if $N \times [0,1]$ were measurable, since $[0,1]$ has positive measure, this would imply that N was measurable too, a contradiction. Hence $A+B$ is not measurable. □

Exercise 15: Consider the function defined over $\mathbb{R}$ by

$$f(x) = \begin{cases} x^{-1/2} & \text{if } 0 < x < 1, \\ 0 & \text{otherwise.} \end{cases}$$

For a fixed enumeration $\{r_n\}$ of the rationals $\mathbb{Q}$, let

$$F(x) = \sum_{n=1}^{\infty} 2^{-n} f(x - r_n).$$

Prove that F is integrable, hence the series defining F converges for almost every $x \in \mathbb{R}$. However, observe that this series is unbounded on every interval, and in fact, any function $\tilde{F}$ that agrees with F a.e. is unbounded in any interval.

Solution. First we compute the integral of f; the improper Riemann integral is

$$\int_0^1 \frac{1}{\sqrt{x}} dx = 2\sqrt{x}\Big|_0^1 = 2,$$

but we only proved that the Lebesgue and Riemann integrals are equal for the proper Riemann integral. Of course it's true for improper integrals as well; here, since $(0,1] = \cup_{n=1}^{\infty}(\frac{1}{n+1}, \frac{1}{n}]$, we have by countable additivity that

$$\begin{aligned} \int_{\mathbb{R}} f dx &= \sum_{n=1}^{\infty} \int_{(\frac{1}{n+1}, \frac{1}{n}]} f dx \\ &= \lim_{N\to\infty} \sum_{n=1}^{N} \int_{(\frac{1}{n+1}, \frac{1}{n}]} f dx \\ &= \lim_{N\to\infty} \int_{\frac{1}{N+1}}^{1} f dx \\ &= \lim_{a\to 0} \int_a^1 f dx \qquad \text{(since this limit exists)} \\ &= 2. \end{aligned}$$

By translation invariance, the integral of $f(x - r_n)$ is also 2. Now since f is nonnegative everywhere, the partial sums are monotonely increasing, so by the Monotone Convergence Theorem

$$\int F dx = \sum_{n=1}^{\infty} \int 2^{-n} f(x - r_n) dx = \sum_{n=1}^{\infty} 2^{1-n} = 2.$$

Since this integral is finite, F is integrable. This implies that F is finite-valued for almost all $x \in \mathbb{R}$.
Now let $\bar{F}$ be any function that agrees with F almost everywhere, and I

any interval on the real line. Let r_N be some rational number contained in I. Then for any $M > 0$, $f(x-r_N) > M$ on the interval $(r_N - \frac{1}{M^2}, r_N + \frac{1}{M^2})$, which intersects I in an interval I_M of positive measure. Since $\bar{F}$ agrees with F almost everywhere, it must also be greater than M at almost all points of this interval $I_M \subset I$. Hence $\bar{F}$ exceeds any finite value M on I. $\square$

Exercise 17: Suppose f is defined on $\mathbb{R}^2$ as follows: $f(x,y) = a_n$ if $n \le x < n+1$ and $n \le y < n+1$, $n \ge 0$; $f(x,y) = -a_n$ if $n \le x < n+1$ and $n+1 \le y < n+2$, $n \ge 0$; $f(x,y) = 0$ elsewhere. Here $a_n = \sum_{k \le n} b_k$, with $\{b_k\}$ a positive sequence such that $\sum_{k=0}^{\infty} b_k = s < \infty$.

(a) Verify that each slice f^y and f_x is integrable. Also for all x, $\int f_x(y)dy = 0$, and hence $\int(\int f(x,y)dy)dx = 0$.

(b) However, $\int f^y(x)dx = a_0$ if $0 \le y < 1$, and $\int f^y(x)dx = a_n - a_{n-1}$ if $n \le y < n+1$ with $n \ge 1$. Hence $y \mapsto \int f^y(x)dx$ is integrable on $(0,\infty)$ and

$$\int \left(\int f(x,y)dx \right) dy = s.$$

(c) Note that $\int_{\mathbb{R}\times\mathbb{R}} |f(x,y)|dxdy = \infty$.

Solution.

(a) Since f is constant on boxes and 0 elsewhere, the horizontal and vertical slices are constant on intervals and 0 elsewhere, and therefore integrable. More precisely,

$$f^y(x) = \begin{cases} -a_{\lfloor y \rfloor - 1} & \lfloor y \rfloor - 1 \le x < \lfloor y \rfloor \\ a_{\lfloor y \rfloor} & \lfloor y \rfloor \le x < \lfloor y \rfloor + 1 \\ 0 & \text{else} \end{cases}$$

for $y \ge 1$,

$$f^y(x) = \begin{cases} a_0 & 0 \le x < 1 \\ 0 & \text{else} \end{cases}$$

for $0 \le y < 1$, and

$$f_x(y) = \begin{cases} a_{\lfloor x \rfloor} & \lfloor x \rfloor \le y < \lfloor x \rfloor + 1 \\ -a_{\lfloor x \rfloor} & \lfloor x \rfloor + 1 \le y < \lfloor x \rfloor + 2 \\ 0 & \text{else} \end{cases}$$

where $\lfloor x \rfloor$ is the greatest integer less than or equal to x. (For $x < 0$ the function $f_x(y)$ is identically 0, and for $y < 0$ the function $f^y(x)$ is identically 0.) Clearly $\int f_x(y)dy = 0$ for all x, since $f_x(y)$ is equal to $a_{\lfloor x \rfloor}$ on an interval of length 1 and $-a_{\lfloor x \rfloor}$ on an interval of length 1 and 0 elsewhere. Hence $\int\int f(x,y)dydx = 0$.

(b) Since all the integrals are of constants on intervals of length 1, it immediately follows from the formulas in part (a) that $\int f^y(x)dx$ is a_0 for $0 \le y < 1$ and $a_n - a_{n-1} = b_n$ for $n \le y < n+1$. Then

$$\int_{\mathbb{R}} \left(\int_{\mathbb{R}} f^y(x)dx \right) = \sum_{n=0}^{\infty} \int_n^{n+1} \left(\int_{\mathbb{R}} f^y(x)dx \right) dy = \sum_{n=0}^{\infty} b_n = s.$$

(c) Since $|f(x,y)|$ is positive, we may use Tonelli's theorem, so

$$\begin{aligned}\int_{\mathbb{R}\times\mathbb{R}} |f(x,y)| &= \int \left(\int f_x(y)dy \right) dx \\ &= \sum_{n=0}^{\infty} \int_n^{n+1} \left(\int f_x(y)dx \right) dx \\ &= \sum_{n=0}^{\infty} 2a_n = \infty\end{aligned}$$

since $a_n > a_0$ so the terms in the sum are bounded away from 0.

□

Exercise 18: Let f be a measurable finite-valued function on $[0,1]$, and suppose that $|f(x) - f(y)|$ is integrable on $[0,1] \times [0,1]$. Show that $f(x)$ is integrable on $[0,1]$.

Solution. Let $g(x,y) = |f(x) - f(y)|$. By Fubini's Theorem, since g is integrable on $[0,1] \times [0,1]$, $g^y(x)$ is an integrable function of x for almost all $y \in [0,1]$. Choose any such y. Then since $f(x) - f(y) \leq |f(x) - f(y)|$,

$$\int_0^1 (f(x) - f(y))\, dx \leq \int_0^1 |f(x) - f(y)| dx < \infty$$

so

$$\int_0^1 f(x)dx \leq f(y) + \int_0^1 |f(x) - f(y)| dx < \infty.$$

□

Exercise 19: Suppose f is integrable on $\mathbb{R}^d$. For each $\alpha > 0$, let $E_\alpha = \{x : |f(x)| > \alpha\}$. Prove that

$$\int_{\mathbb{R}^d} |f(x)| dx = \int_0^\infty m(E_\alpha) d\alpha.$$

Solution. By Tonelli's Theorem,

$$\begin{aligned}\int_0^\infty m(E_\alpha) d\alpha &= \int_0^\infty \left(\int_{\mathbb{R}^d} \chi_{|f(x)|>\alpha} dx \right) d\alpha \\ &= \int_{\mathbb{R}^d} \left(\int_0^\infty \chi_{|f(x)|>\alpha} d\alpha \right) dx \\ &= \int_{\mathbb{R}^d} |f(x)| dx.\end{aligned}$$

□

Exercise 22: Prove that if $f \in L^1(\mathbb{R}^d)$ and

$$\hat{f}(\xi) = \int_{\mathbb{R}^d} f(x) e^{-2\pi i x \xi} dx,$$

then $\hat{f}(\xi) \to 0$ as $|\xi| \to \infty$. (This is the Riemann-Lebesgue lemma.)

Solution. By translation invariance,

$$\begin{aligned}
\hat{f}(\xi) &= \int_{\mathbb{R}^d} f(x)e^{-2\pi i x\cdot\xi}dx \\
&= \int_{\mathbb{R}^d} f\left(x - \frac{\xi}{2|\xi|^2}\right) e^{-2\pi i (x - \frac{\xi}{2|\xi|^2})\cdot\xi}dx \\
&= \int_{\mathbb{R}^d} f\left(x - \frac{\xi}{|\xi|^2}\right) e^{-2\pi i x\cdot\xi} e^{2\pi i \frac{\xi}{2|\xi|^2}\cdot\xi}dx \\
&= \int_{\mathbb{R}^d} -f\left(x - \frac{\xi}{2|\xi|^2}\right) e^{-2\pi i x\cdot\xi}dx
\end{aligned}$$

so, multiplying by $\frac{1}{2}$ and adding the original expression,

$$\begin{aligned}
\hat{f}(\xi) &= \frac{1}{2}\left(\int_{\mathbb{R}^d} f(x)e^{-2\pi i x\cdot\xi}dx - \int_{\mathbb{R}^d} f\left(x - \frac{\xi}{2|\xi|^2}\right) e^{-2\pi i x\cdot\xi}dx\right) \\
&= \frac{1}{2}\int_{\mathbb{R}^d}\left(f(x) - f\left(x - \frac{\xi}{2|\xi|^2}\right)\right) e^{-2\pi i x\cdot\xi}dx
\end{aligned}$$

and

$$\begin{aligned}
|\hat{f}(\xi)| &= \frac{1}{2}\left|\int_{\mathbb{R}^d}\left(f(x) - f\left(x - \frac{\xi}{2|\xi|^2}\right)\right) e^{-2\pi i x\cdot\xi}dx\right| \\
&\leq \frac{1}{2}\int_{\mathbb{R}^d}\left|\left(f(x) - f\left(x - \frac{\xi}{2|\xi|^2}\right)\right) e^{-2\pi i x\cdot\xi}\right| dx \\
&= \frac{1}{2}\int_{\mathbb{R}^d}\left|f(x) - f\left(x - \frac{\xi}{2|\xi|^2}\right)\right| dx \\
&= \frac{1}{2}\left\|f(x) - f\left(x - \frac{\xi}{2|\xi|^2}\right)\right\|_1 .
\end{aligned}$$

As $|\xi| \to \infty$, $|\frac{\xi}{2|\xi|^2}| \to 0$, so $\|f(x) - f(x - \frac{\xi}{2|\xi|^2})\|_1 \to 0$ by the L^1-continuity of translation (Proposition 2.5). □

Exercise 23: As an application of the Fourier transform, show that there does not exist a function $I \in L^1(\mathbb{R}^d)$ such that

$$f * I = f \quad \text{for all } f \in L^1(\mathbb{R}^d).$$

Solution. Suppose such an I exists. Then for every $f \in L^1$, $\hat{f}(\xi)\hat{I}(\xi) = \hat{f}(\xi)$ for all ξ. This implies that $\hat{I}(\xi) = 1$ for all ξ. But this contradicts the Riemann-Lebesgue Lemma (Problem 22). QED.
Note: To be totally complete, we should show that for any ξ there is a function $g \in L^1$ such that $\hat{g}(\xi) \neq 0$. Otherwise, the equation $\hat{f}\hat{I} = \hat{f}$ wouldn't necessarily imply $\hat{I} = 1$ everywhere. But it is easy to show that such a g exists for any ξ; for example, g could be equal to $e^{2\pi i x\cdot\xi}$ on some compact set and 0 outside. □

Exercise 24: Consider the convolution

$$(f * g)(x) = \int_{\mathbb{R}^d} f(x - y)g(y)dy.$$

(a) Show that $f * g$ is uniformly continuous when f is integrable and g bounded.

(b) If in addition g is integrable, prove that $(f * g)(x) \to 0$ as $|x| \to \infty$.

Solution.

(a) Since g is bounded, $\exists M$ with $|g| < M$ everywhere. Then

$$\begin{aligned}
|f * g(x) - f * g(x')| &= \left| \int_{\mathbb{R}^d} f(x-y)g(y)dy - \int_{\mathbb{R}^d} f(x'-y)g(y)dy \right| \\
&= \left| \int_{\mathbb{R}^d} \left(f(x-y) - f(x'-y)\right) g(y)dy \right| \\
&\leq \int_{\mathbb{R}^d} |f(x-y) - f(x'-y)|\, |g(y)|dy \\
&\leq M \int_{\mathbb{R}^d} |f(x-y) - f(x'-y)|dy \\
&= M \int_{\mathbb{R}^d} |f(y) - f(y + (x'-x))|dy \\
&= M\|f(y) - f(y + (x-x'))\|_1.
\end{aligned}$$

In the penultimate step we have used translation invariance and the fact that $\int f(y)dy = \int f(-y)dy$ provided both integrals are taken over all of $\mathbb{R}^d$. Now by the L^1-continuity of translation, $\exists \delta > 0$ such that $|x - x'| < \delta \Rightarrow \|f(y) - f(y + (x-x'))\|_1 < \frac{\epsilon}{M}$. This in turn implies $|f * g(x) - f * g(x')| < \epsilon$, so $f * g$ is uniformly continuous.

(b) Let $\epsilon > 0$. Since $C_C(\mathbb{R}^d)$ is dense in $L^1(\mathbb{R}^d)$, we may choose $\tilde{f}$ such that $\operatorname{supp}(\tilde{f}) = K$ is compact, $\tilde{f}$ is continuous, and $\|f - \tilde{f}\|_1 < \frac{\epsilon}{2M}$, where M is a bound for $|g|$ as in part (a). Continuous functions with compact support are bounded, so choose N such that $|\tilde{f}| < N$. Now

$$\begin{aligned}
|f * g(x)| &= \left| \int_{\mathbb{R}^d} f(x-y)g(y)dy \right| \\
&= \left| \int_{\mathbb{R}^d} \left(\tilde{f}(x-y) + (f(x-y) - \tilde{f}(x-y))\right) g(y)dy \right| \\
&\leq \int_{\mathbb{R}^d} \left| \tilde{f}(x-y) + (f(x-y) - \tilde{f}(x-y))g(y) \right| dy \\
&\leq \int_{\mathbb{R}^d} |\tilde{f}(x-y)||g(y)|dy + \int_{\mathbb{R}^d} \left| f(x-y) - \tilde{f}(x-y) \right| |g(y)|dy.
\end{aligned}$$

Call the first integral I_1 and the second I_2. Since $|g| < M$, $I_2 \leq M\|f - \tilde{f}\|_1 < \frac{\epsilon}{2}$. Now since g is integrable, there must exist compact F such that $\int_{F^c} |g| < \frac{\epsilon}{2N}$. Then if $|x|$ is larger than the sum of the diameters of K and N,

$$\begin{aligned}
I_2 &= \int_{\mathbb{R}^d} |\tilde{f}(x-y)||g(y)|dy \\
&= \int_{x-y \in K} |\tilde{f}(x-y)||g(y)|dy \qquad \text{since } \tilde{f} = 0 \text{ on } K^c \\
&\leq N \int_{x-y \in K} |g(y)|dy \\
&\leq N \int_{F^c} |g(y)|dy < \frac{\epsilon}{2}
\end{aligned}$$

since $\{y : x - y \in K\} \subset F^c$. Thus, for sufficiently large x, $|f * g(x)| \leq I_1 + I_2 < \epsilon$.

□

Chapter 2.6, Page 95

Problem 1: If f is integrable on $[0, 2\pi]$, then $\int_0^{2\pi} f(x)e^{-inx}dx \to 0$ as $|n| \to \infty$. Show as a consequence that if E is a measurable subset of $[0, 2\pi]$, then

$$\int_E \cos^2(nx + u_n)dx \to \frac{m(E)}{2}, \quad \text{as } n \to \infty$$

for any sequence $\{u_n\}$.

Solution. First, note that $\int_0^{2\pi} f(x)\cos(nx)dx \to 0$ and $\int_0^{2\pi} f(x)\sin(nx) \to 0$ since these are the real and imaginary parts of $\int_0^{2\pi} f(x)e^{-inx}dx$. In particular, if we let $f(x) = \chi_E(x)$ for some measurable $E \subset [0, 2\pi]$, then for any $\epsilon > 0$, $\exists N$ such that $|\int_0^{2\pi} \chi_E(x)\sin(nx)dx|$ and $|\int_0^{2\pi} \chi_E(x)\cos(nx)dx|$ are both less than $\frac{\epsilon}{2}$ provided $|n| > N$. Then for any sequence u_n,

$$\begin{aligned}
\left|\int_E \cos(2nx + 2u_n)dx\right| &= \left|\int_E \cos(2nx)\cos(2u_n) - \sin(2nx)\sin(2u_n)dx\right| \\
&= \left|\cos(2u_n)\int_0^{2\pi} \chi_E(x)\cos(2nx)dx - \sin(2u_n)\int_0^{2\pi} \chi_E(x)\sin(2nx)dx\right| \\
&\leq |\cos(2u_n)|\left|\int_0^{2\pi} \chi_E(x)\cos(2nx)dx\right| + |\sin(2u_n)|\left|\int_0^{2\pi} \chi_E(x)\sin(2nx)dx\right| \\
&\leq 1 \cdot \frac{\epsilon}{2} + 1 \cdot \frac{\epsilon}{2} = \epsilon
\end{aligned}$$

for $|n| > N$. Hence $\int_E \cos(2nx + u_n)dx \to 0$ as $|n| \to \infty$. Now

$$\begin{aligned}
\int_E \cos^2(nx + u_n) &= \int_E \frac{1}{2}\left(1 + \cos(2(nx + u_n))\right)dx \\
&= \frac{m(E)}{2} + \int_0^{2\pi} \chi_E(x)\cos(2nx + 2u_n)dx
\end{aligned}$$

and we have shown that the second term tends to 0 as $|n| \to \infty$. □

Problem 2: Prove the Cantor-Lebesgue theorem: if

$$\sum_{n=0}^{\infty} A_n(x) = \sum_{n=0}^{\infty}(a_n \cos nx + b_n \sin nx)$$

converges for x in a set of positive measure (or in particular for all x), then $a_n \to 0$ and $b_n \to 0$ as $n \to \infty$.

Solution. We can rewrite $A_n(x) = c_n \cos(nx + d_n)$ where $c_n = \sqrt{a_n^2 + b_n^2}$ and d_n is some phase angle (it can be $-\arctan(b_n/a_n)$, for example). If $\sum A_n(x)$ converges on some E with $m(E) > 0$, then $A_n(x) \to 0$ on E. By

Egorov's theorem, this implies $A_n \to 0$ uniformly on some $E' \subset e$ with $m(E') > 0$. Then

$$\begin{aligned} c_n \cos(nx + d_n) &\xrightarrow{u} 0 \qquad \text{on } E' \\ \Rightarrow c_n^2 \cos^2(nx + d_n) &\xrightarrow{u} 0 \qquad \text{on } E' \\ \Rightarrow \int_{E'} c_n^2 \cos^2(nx + d_n)dx &\to 0. \end{aligned}$$

But $\int_E' \cos^2(nx+d_n)dx \to \frac{m(E')}{2}$ by the previous problem, so $c_n^2 \to 0$, which implies $c_n \to 0$, which implies $a_n \to 0$ and $b_n \to 0$. □

Problem 3: A sequence $\{f_k\}$ of measurable functions on $\mathbb{R}^d$ is **Cauchy in measure** if for every $\epsilon > 0$,

$$m(\{x : |f_k(x) - f_\ell(x)| > \epsilon\}) \to 0 \quad \text{as } k, \ell \to \infty.$$

We say that $\{f_k\}$ **converges in measure** to a (measurable) function f if for every $\epsilon > 0$,

$$m(\{x : |f_k(x) - f(x)| > \epsilon\}) \to 0 \quad \text{as } k \to \infty.$$

This notion coincides with the "convergence in probability" of probability theory.
Prove that if a sequence $\{f_k\}$ of integrable functions converges to f in L^1, then $\{f_k\}$ converges to f in measure. Is the converse true?

Solution. Suppose $f_n, f \in L^1$ and $f_n \to f$ in L^1. By Chebyshev's Inequality,

$$m\left(\{x : |f_n(x) - f(x)| > \epsilon\}\right) \leq \frac{\|f_n - f\|_1}{\epsilon}$$

and since the RHS tends to 0 as $n \to \infty$, the LHS does as well, so $f_n \to f$ in measure. However, the converse does not hold. Consider the case $f(x) = 0, f_n(x) = n\chi_{[0,\frac{1}{n}]}$. Then $m(\{x : f(x) \neq f_n(x)\}) = \frac{1}{n}$ so for any ϵ, $m(\{x : |f(x) - f_n(x)| > \epsilon\}) \leq \frac{1}{n} \to 0$. Hence $f_n \to f$ in measure. However, $\|f_n - f\|_1 = 1$ for all n, so $f_n \not\to f$ in L^1. □

Problem 4: We have already seen (in Exercise 8, Chapter 1) that if E is a measurable set in $\mathbb{R}^d$, and L is a linear transformation of $\mathbb{R}^d$ to $\mathbb{R}^d$, then $L(E)$ is also measurable, and if E has measure 0, then so has $L(E)$. The quantitative statement is

$$m(L(E)) = |\det(L)|m(E).$$

As a special case, note that the Lebesgue measure is invariant under rotations. (For this special case see also Exercise 26 in the next chapter.)
The above identity can be proved using Fubini's theorem as follows.

(a) Consider first the case $d = 2$, and L a "strictly" upper triangular transformation $x' = x + ay, y' = y$. Then

$$\chi_{L(E)}(x, y) = \chi_E(L^{-1}(x, y)) = \chi_E(x - ay, y).$$

Hence

$$\begin{aligned} m(L(E)) &= \int_{\mathbb{R}\times\mathbb{R}} \left(\int \chi_E(x - ay, y)dx \right) dy \\ &= \int_{\mathbb{R}\times\mathbb{R}} \left(\int \chi_E(x, y)dx \right) dy \\ &= m(E), \end{aligned}$$

by the translation-invariance of the measure.

(b) Similarly $m(L(E)) = m(E)$ if L is strictly lower triangular. In general, one can write $L = L_1 \Delta L_2$, where L_j are strictly (upper and lower) triangular and Δ is diagonal. Thus $m(L(E)) = |\det L| m(E)$, if one uses Exercise 7 in Chapter 1.

Solution.

(a) I'm not quite sure what to comment on here, since the problem statement pretty much did all the work for me. I guess I should point out that the use of Tonelli's theorem to turn the double integral into an iterated integral is justified because χ_E is nonnegative; note that this proves that the result holds even if $m(E) = \infty$.

(b) The fact that lower triangular transformations work the same way is obvious. Now supposing $L = L_1 \Delta L_2$, we have

$$\begin{aligned} m(L(E)) &= m(L_1(\Delta(L_2(E)))) = m(\Delta(L_2(E))) \\ &= |\det(\Delta)| m(L_2(E)) = |\det(\Delta)| m(E) = |\det(L)| m(E) \end{aligned}$$

since $\det(L) = \det(L_1)\det(\Delta)\det(L_2) = \det(\Delta)$ and $m(\Delta(E)) = |\det(\Delta)| m(E)$ by Exercise 7 of Chapter 1. Thus, every linear transformation that has an LU decomposition works as we want it to. However, I think the problem is flawed because not every matrix has an LU decomposition (in fact, not even every *invertible* matrix does.) In particular, in the 2×2 case a matrix of the form $L_1 \Delta L_2$ will look like either

$$\begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix} \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ e & 1 \end{pmatrix} = \begin{pmatrix} d_1 + cd_2 e & cd_2 \\ d_2 e & d_2 \end{pmatrix}$$

or

$$\begin{pmatrix} 1 & 0 \\ e & 1 \end{pmatrix} \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} \begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} d_1 & cd_1 \\ ed_1 & ed_1 c + d_2 \end{pmatrix}.$$

But a matrix of the form

$$\begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix}$$

with $a, b \neq 0$ cannot be put in either form, because in the first case we would need $d_2 = 0$ in order to make the lower right entry 0, and then the upper right and lower left entries could not be nonzero; similarly, in the second case, we would need $d_1 = 0$ which would make $a, b \neq 0$ impossible. Hence, there are some matrices that cannot be factored in the way this problem indicates; for the ones that can, though, we know that they expand measures by a factor of $|\det|$.

□

Chapter 3.5, Page 145

Exercise 1: Suppose ϕ is an integrable function on $\mathbb{R}^d$ with $\int_{\mathbb{R}^d} \phi(x)dx = 1$. Set $K_\delta(x) = \delta^{-d}\phi(x/\delta)$, $\delta > 0$.

(a) Prove that $\{K_\delta\}_{\delta>0}$ is a family of good kernels.

(b) Assume in addition that ϕ is bounded and supported in a bounded set. Verify that $\{K_\delta\}_{\delta>0}$ is an approximation to the identity.

(c) Show that Theorem 2.3 holds for good kernels as well.

Solution.

(a) By the dilation properties of the integral, we have immediately that $\int_{\mathbb{R}^d} K_\delta(x)dx = \int_{\mathbb{R}^d} \phi(x)dx = 1$ and $\int_{\mathbb{R}^d} |K_\delta(x)|dx = \int_{\mathbb{R}^d} |\phi(x)|dx = \|\phi\|_1 < \infty$. This proves the first two properties of good kernels. For the last, we recall that for $\phi \in L^1$, for every $\epsilon > 0$ there exists a compact set F_ϵ such that $\int_{F_\epsilon^c} |\phi| < \epsilon$. Now compact subsets of $\mathbb{R}^d$ are bounded, so $K_\epsilon \subset B_{r_\epsilon}(0)$ for some radius r_ϵ. Now if $h > 0$ is any fixed number, for $\delta < \frac{h}{r_\epsilon}$ this will imply that $\int_{|x|>h} |K_\delta(x)|dx < \epsilon$. Thus, for any $h > 0$, $\int_{|x|>h} |K_\delta(x)|dx \to 0$ as $\delta \to 0$. Hence $\{K_\delta\}$ is a family of good kernels.

(b) Suppose $|\phi| \leq M$ everywhere and $\phi(x) = 0$ for $|x| \geq B$. Let $A = MB^{d+1}$. Then for any $\delta > 0$,

$$\frac{1}{\delta}|K_\delta(x)| = \delta^{-(d+1)}\phi\left(\frac{x}{\delta}\right) \leq M\delta^{-(d+1)}$$
$$\leq \frac{A}{(\delta B)^{d+1}} \leq \frac{A}{|x|^{d+1}}$$

for $\frac{x}{\delta} \leq B$; for $\frac{x}{\delta} > B$ we have $K_\delta(x) = 0$. Hence $|K_\delta(x)| \leq A\delta/|x|^{d+1}$ for all x and δ, so $\{K_\delta\}$ is an approximation to the identity.

(c) Suppose $\{K_\delta\}$ is any family of good kernels. Then

$$\begin{aligned}
\|f * K_\delta - f\| &= \int_{\mathbb{R}^d} |f * K_\delta(x) - f(x)|\,dx \\
&= \int_{\mathbb{R}^d} \left|\int_{\mathbb{R}^d} f(x-y)K_\delta(y)dy - f(x)\right| dx \\
&= \int_{\mathbb{R}^d} \left|\int_{\mathbb{R}^d} (f(x-y) - f(x))K_\delta(y)dy\right| dx \\
&\leq \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |f(x-y) - f(x)||K_\delta(y)|dydx \\
&= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |f(x-y) - f(x)||K_\delta(y)|dxdy \\
&= \int_{|y|\leq\eta}\int_{\mathbb{R}^d} |f(x-y) - f(x)||K_\delta(y)|dxdy + \int_{|y|>\eta}\int_{\mathbb{R}^d} |f(x-y) - f(x)||K_\delta(y)|dxdy
\end{aligned}$$

for any $\eta > 0$. Let us call the first integral I_1 and the second I_2. Then

$$\begin{aligned}
I_1 &= \int_{|y|\leq\eta} |K_\delta(y)| \left(\int_{\mathbb{R}^d} |f(x-y) - f(x)|dx\right) dy \\
&= \int_{|y|\leq\eta} |K_\delta(y)|\|f(x-y) - f(x)\|_1 dy.
\end{aligned}$$

Now by the L^1-continuity of translation, $\|f(x-y)-f(x)\|_1 \to 0$ as $y \to 0$. Thus, if η is sufficiently small, this norm will be at most, say, $\frac{\epsilon}{2A}$ for all $y \in [-\eta, \eta]$. Then

$$I_1 \le \int_{|y|\le\eta} |K_\delta(y)| \frac{\epsilon}{2A} dy \le \frac{\epsilon}{2}$$

since $\int_{\mathbb{R}^d} |K_\delta(y)|dy \le A$ for all δ. Thus, by choosing η sufficiently small, we may make I_1 as small as we like, independent of δ. On the other hand,

$$\begin{aligned} I_2 &= \int_{|y|>\eta} |K_\delta(y)| \int_{\mathbb{R}^d} |f(x-y)-f(x)|dxdy \\ &\le \int_{|y|>\eta} |K_\delta(y)| \int_{\mathbb{R}^d} (|f(x-y)|+|f(x)|)dxdy \\ &= \int_{|y|>\eta} |K_\delta(y)| \cdot 2\|f\|_1 dy \\ &= 2\|f\|_1 \int_{|y|>\eta} |K_\delta(y)|dy \to 0 \end{aligned}$$

as $\delta \to 0$. Putting the two halves together, we see that choosing a sufficiently small η and then letting $\delta \to 0$ makes $\|f * K_\delta - f\|_1 \to 0$. □

Exercise 3: Suppose 0 is a point of (Lebesgue) density of the set $E \subset \mathbb{R}$. Show that for each of the individual conditions below there is an infinite sequence of points $x_n \in E$, with $x_n \neq 0$, and $x_n \to 0$ as $n \to \infty$.
(a) The sequence also satisfies $-x_n \in E$ for all n.
(b) In addition, $2x_n$ belongs to E for all n.
Generalize.

Solution.

(a) Since 0 is a point of density of E, there exists $r_0 > 0$ such that $m(E \cap B_r(0)) > \frac{2}{3}m(B_r(0)) = \frac{4}{3}r$ for $r \le r_0$. By symmetry, $m((-E) \cap B_{r_0}(0)) = m(E \cap B_{r_0}(0)) \ge \frac{4}{3}r_0$. Now E and $-E$ both intersect $B_{r_0}(0)$, which has measure $2r_0$, in sets of measure at least $\frac{4}{3}r_0$. Therefore they intersect each other in a set of measure at least $\frac{2}{3}r_0$. Since $E \cap (-E)$ has positive measure, it is infinite, so it contains an infinite sequence x_n. This sequence satisfies $x_n \in E$ and $-x_n \in E$.

(b) Since 0 is a point of density of E, there exists $r_0 > 0$ such that $m(E \cap B_r(0)) > \frac{2}{3}m(B_r(0)) = \frac{4}{3}r$ for $r \le r_0$. Let $E_0 = E \cap B_{r_0}(0)$. Then $m(\frac{1}{2}E_0) = \frac{1}{2}m(E_0) \ge \frac{2}{3}r_0$ as we showed in a previous homework about the effect of dilation on Lebesgue measure. Now $\frac{1}{2}E_0 = (\frac{1}{2}E) \cap B_{r_0/2}(0)$ has measure at least $\frac{2}{3}r_0$, and we also know $m(E \cap B_{r_0/2}(0)) \ge \frac{2}{3}m(B_{r_0/2}) = \frac{2}{3}r_0$ since $\frac{r_0}{2} < r_0$. So E and $\frac{1}{2}E$ both intersect $B_{r_0/2}$, which has measure r_0, in sets of measure at least $\frac{2}{3}r_0$. Therefore they intersect each other in a set of measure at least $\frac{1}{3}r_0$. Since $E \cap (\frac{1}{2}E)$ has positive measure, it must contain an infinite sequence x_n. Then $x_n \in E$ and $2x_n \in E$.

Clearly the above process generalizes to produce a sequence x_n with $x_n \in E$ and $cx_n \in E$ for any $c \neq 0$. $\square$

Exercise 5: Consider the function on $\mathbb{R}$ defined by

$$f(x) = \begin{cases} \frac{1}{|x|(\log 1/|x|)^2} & \text{if } |x| \leq 1/2, \\ 0 & \text{otherwise.} \end{cases}$$

(a) Verify that f is integrable.
(b) Establish the inequality

$$f^*(x) \geq \frac{c}{|x|(\log 1/|x|)} \qquad \text{for some } c > 0 \text{ and all } |x| \leq 1/2,$$

to conclude that the maximal function f^* is not locally integrable.

Solution.
(a) We have

$$\int_{\mathbb{R}^d} |f(x)|dx = 2\int_0^{1/2} \frac{1}{x(\log 1/x)^2}dx = 2\frac{1}{\log \frac{1}{x}}\Bigg|_0^{1/2} = \frac{2}{\log 2}.$$

(b) For $0 < |x| \leq \frac{1}{2}$, if $B = (0, 2x)$ is the ball of radius $|x|$ centered at x, then

$$\begin{aligned} \frac{1}{m(B)}\int_B |f(y)|dy &= \frac{1}{2|x|}\int_0^2 |x||f(y)|dy \\ &\geq \frac{1}{2|x|}\int_0^| x|\frac{1}{y(\log 1/y)^2}dy \\ &= \frac{1}{2|x|}\frac{1}{\log(1/|x|)}. \end{aligned}$$

Since $f^*(x)$ is the supremum of such integrals over all balls containing x, it is at least equal to the integral over B, so $f^*(x) \geq \frac{1}{2|x|\log(1/|x|)}$. This function is not locally integrable, because if we integrate it in any neighborhood around 0 we get

$$\int_{-\delta}^{\delta} \frac{1}{2|x|\log(1/x)}dx = -\log\left(\log\frac{1}{x}\right)\Bigg|_0^{\delta} = \infty.$$

$\square$

Exercise 6: In one dimension there is a version of the basic inequality (1) for the maximal function in the form of an identity. We defined the "one-sided" maximal function

$$f_+^*(x) = \sup_{h>0}\frac{1}{h}\int_x^{x+h} |f(y)|dx.$$

If $E_\alpha^+ = \{x \in \mathbb{R} : f_+^*(x) > \alpha\}$, then

$$m(E_\alpha^+) = \frac{1}{\alpha}\int_{E_\alpha^+} |f(y)|dy.$$

Solution. First, we note that

$$\begin{aligned} x \in E_\alpha^+ &\Leftrightarrow \exists h > 0 \text{ s.t. } \frac{1}{h}\int_x^{x+h} |f(y)|dy > \alpha \\ &\Leftrightarrow \int_x^{x+h} |f(y)|dy > \alpha h \\ &\Leftrightarrow \int_0^{x+h} |f(y)|dy - \int_0^x |f(y)|dy - \alpha(x+h) + \alpha x > 0 \\ &\Leftrightarrow \int_0^{x+h} |f(y)|dy - \alpha(x+h) > \int_0^x |f(y)|dy - \alpha x. \end{aligned}$$

Thus, if we define $F(x) = \int_0^x |f(y)|dy - \alpha x$, the set E_α^+ is precisely the set $\{x : \exists h > 0 \text{ s.t. } F(x+h) > F(x)\}$. Note also that F is continuous by the absolute continuity of the integral (we are assuming that $f \in L^1$, naturally). By the Rising Sun Lemma,

$$E_\alpha^+ = \bigcup_{j=1}^{\infty} (a_j, b_j)$$

where the intervals (a_j, b_j) are disjoint and

$$\begin{aligned} F(a_j) &= F(b_j) \\ \Rightarrow \int_0^{a_j} |f(y)|dy - \alpha a_j &= \int_0^{b_j} |f(y)|dy - \alpha b_j \\ \Rightarrow \int_{a_j}^{b_j} |f(y)|dy &= \alpha(b_j - a_j). \end{aligned}$$

Then

$$\int_{E_\alpha^+} |f(y)|dy = \sum_{j=1}^{\infty} \int_{a_j}^{b_j} |f(y)|dy = \alpha \sum_{j=1}^{\infty} (b_j - a_j) = \alpha m(E_\alpha^+)$$

as desired. □

Exercise 8: Suppose A is a Lebesgue measurable set in $\mathbb{R}$ with $m(A) > 0$. Does there exist a sequence $\{s_n\}$ such that the complement of $\cup_{n=1}^{\infty}(A+s_n)$ in $\mathbb{R}$ has measure zero?

Solution. Yes. Let x be any point of density of A, and let $s_n = q_n - x$ where $\{q_n\}$ is an enumeration of the rationals. Let $E = \cup(A + s_n)$. To show that $m(E^c) = 0$, it is sufficient to show that $m(E^c \cap [n, n+1]) = 0$ for all $n \in \mathbb{Z}$; since E is invariant under rational translations, it is sufficient to show that $E^c \cap [0,1]$ has measure zero.
For $m = 1, 2, \ldots$ let N_m be an integer such that $m(A \cap B_{1/N_m}(x)) \geq (1 - \frac{1}{m})m(B_{1/N_m})(x)$. Such an N_m must exist because x is a point of density of A. Then by the construction of E, $m(A \cap B_{1/N_m}(q)) \geq (1 - \frac{1}{m})m(B_{1/N_m})(q)$ for any $q \in \mathbb{Q}$. Now the open balls

$$U_m^j = B_{1/N_m}\left(\frac{j}{2N_m}\right), \; j = 1, 2, \ldots, 2N_m$$

cover $[0,1]$, so

$$E^c \cap [0,1] \subset \bigcup_{j=1}^{2N_m} \left(E^c \cap U_m^j\right)$$

and

$$m\left(E^c \cap [0,1]\right) \leq \sum_{j=1}^{2N_m} m\left(E^c \cap U_m^j\right) \leq \sum_{j=1}^{2N_m} \frac{1}{m} m(U_m^j) = (2N_m)\frac{1}{m}\frac{2}{N_m} = \frac{4}{m}.$$

Since $m(E^c \cap [0,1]) \leq \frac{4}{m}$ for all m, $m(E^c \cap [0,1]) = 0$. Hence $m(E^c) = 0$. $\square$

Note: It is not sufficient to construct a set whose Lebesgue points are dense in $\mathbb{R}$. Consider an open dense set of measure $\leq \epsilon$ (e.g. put an interval of length $\frac{\epsilon}{2^k}$ around the kth rational). Then every point is a Lebesgue point since the set is open, yet its complement has positive measure.

Exercise 9: Let F be a closed subset in $\mathbb{R}$, and $\delta(x)$ the distance function from x to F, that is

$$\delta(x) = \inf\{|x-y| : y \in F\}.$$

Clearly, $\delta(x+y) \leq |y|$ whenever $x \in F$. Prove the more refined estimate

$$\delta(x+y) = o(|y|) \qquad \text{for a.e. } x \in F,$$

that is, $\delta(x+y)/|y| \to 0$ for a.e. $x \in F$.

Solution. We note that δ is a function of bounded variation on any interval; in fact, in general we have $V_a^b(\delta) \leq |b-a|$ because $|\delta(y) - \delta(z)| \leq |y-z|$ for all $y, z \in \mathbb{R}$. Since δ has bounded variation, it is differentiable almost everywhere; in particular, it is differentiable for a.e. $x \in F$. But δ is a nonnegative function that is 0 for all $x \in F$, so it has a local minimum at every point of F; if it is differentiable at $x \in F$, its derivative is zero. By the definition of derivative, this implies $\delta(x+y)/|y| \to 0$. $\square$

Exercise 15: Suppose F is of bounded variation and continuous. Prove that $F = F_1 - F_2$, where both F_1 and F_2 are monotonic and continuous.

Solution. Every bounded-variation function is a difference of increasing functions, so write $F = G_1 - G_2$ where G_1 and G_2 are increasing. As shown in Lemmas 3.12-13, an increasing function is a continuous increasing function plus a jump function. Hence $G_1 = F_1 + J_1$ where F_1 is continuous and increasing, and J_1 is a jump function; similarly, $G_2 = F_2 + J_2$. Then $F = (F_1 - F_2) + (J_1 - J_2)$. But $J_1 - J_2$ is a jump function, and jump functions are continuous only if they're constant. Since F is continuous, this implies that $J_1 - J_2$ is constant; WLOG, $J_1 - J_2 = 0$. (Otherwise we could redefine $F_1' = F_1 + (J_1 - J_2)$ and F_1' would also be continuous and increasing.) Hence $F = F_1 - F_2$. $\square$

Exercise 17: Prove that if $\{K_\epsilon\}_{\epsilon>0}$ is a family of approximations to the identity, then

$$\sup_{\epsilon>0} |(f * K_\epsilon)(x)| \leq c f^*(x)$$

for some constant $c > 0$ and all integrable f.

Exercise 18: Verify the agreement between the two definitions given for the Cantor-Lebesgue function in Exercise 2, Chapter 1 and Section 3.1 of this chapter.

Solution. This is such a lame problem. It's so clear that they're the same. Probably the easiest way to see that is to think of the Cantor-Lebesgue function as the following process:

- Given x, let y be the greatest member of the Cantor set such that $y \leq x$. (We know such a y exists because the Cantor set is closed.)
- Write the ternary expansion of y.
- Change all the 2's to 1's and re-interpret as a binary expansion. The value obtained is $F(x)$.

It's pretty clear that both the definitions of the Cantor-Lebesgue function given in the text do exactly this. □

Exercise 19: Show that if $f : \mathbb{R} \to \mathbb{R}$ is absolutely continuous, then

(a) f maps sets of measure zero to sets of measure zero.
(b) f maps measurable sets to measurable sets.

Solution.

(a) Suppose $E \subset \mathbb{R}$ has measure zero. Let $\epsilon > 0$. By absolute continuity, $\exists \delta > 0$ such that $\sum |b_j - a_j| < \delta \Rightarrow \sum |f(b_j) - f(a_j)| < \epsilon$. Since $m(E) = 0$, there is an open set $U \supset E$ with $m(U) < \delta$. Every open subset of $\mathbb{R}$ is a countable disjoint union of open intervals, so

$$U = \bigcup_{j=1}^{\infty} (a_j, b_j) \text{ with } \sum_{j=1}^{\infty} (b_j - a_j) < \delta.$$

For each $j = 1, 2, \ldots,$, let $m_j, M_j \in [a_j, b_j]$ be values of x with

$$f(m_j) = \min_{x \in [a,b]} f(x) \text{ and } f(M_j) = \max_{x \in [a,b]} f(x).$$

Such m_j and M_j must exist because f is continuous and $[a_j, b_j]$ is compact. Then

$$f(U) \subset \bigcup_{j=1}^{\infty} [f(m_j), f(M_j)].$$

But $|M_j - m_j| \leq |b_j - a_j|$ so

$$\sum_{j=1}^{\infty} |M_j - m_j| < \delta \Rightarrow \sum_{j=1}^{\infty} |f(M_j) - f(m_j)| < \epsilon.$$

Hence $f(E)$ is a subset of a set of measure less that ϵ. This is true for all ϵ, so $f(E)$ has measure zero.

(b) Let $E \subset \mathbb{R}$ be measurable. Then $E = F \cup N$ where F is F_σ and N has measure zero. Since closed subsets of $\mathbb{R}$ are σ-compact, F is σ-compact. But then $f(F)$ is also σ-compact since f is continuous. Then $f(E) = f(F) \cup f(N)$ is a union of an F_σ set and a set of measure zero. Hence $f(E)$ is measurable.

□

Exercise 20: This exercise deals with functions F that are absolutely continuous on $[a,b]$ and are increasing. Let $A = F(a)$ and $B = F(b)$.

(a) There exists such an F that is in addition strictly increasing, but such that $F'(x) = 0$ on a set of positive measure.

(b) The F in (a) can be chosen so that there is a measurable subset $E \subset [A,B]$, $m(E) = 0$, so that $F^{-1}(E)$ is not measurable.

(c) Prove, however, that for any increasing absolutely continuous F, and E a measurable subset of $[A,B]$, the set $F^{-1}(E) \cap \{F'(x) > 0\}$ is measurable.

Solution.

(a) Let

$$F(x) = \int_a^x \delta_C(x)dx$$

where $C \subset [a,b]$ is a Cantor set of positive measure and $\delta_C(x)$ is the distance from x to C. Note that $\delta_C(x) \geq 0$ with equality iff $x \in C$. Since δ_C is continuous, this integral is well-defined, even in the Riemann sense. Moreover, F is absolutely continuous by the absolute continuity of integration of L^1 functions. As shown in problem 9, $F'(x)$ exists and equals zero a.e. in C, hence on a set of positive measure. However, F is strictly increasing: Suppose $a \leq x < y \leq b$. Since C contains no interval, some point, and therefore some interval, between x and y belongs to C^C. The integral of δ_C over this interval will be positive, so $F(y) > F(x)$.

(b) The same function from part (a) does the trick. Since F is increasing, it maps disjoint open intervals to disjoint open intervals. Let $U = [a,b] \setminus C$. Since U is open, we can write

$$U = \bigcup_{j=1}^{\infty} (a_j, b_j)$$

where the intervals (a_j, b_j) are disjoint. Then

$$F(U) = \bigcup_{j=1}^{\infty} (F(a_j), F(b_j))$$

and

$$m(F(U)) = \sum_{j=1}^{\infty} F((b_j) - F(a_j)).$$

But

$$B - A = F(b) - F(a) = \int_a^b \delta(x)dx = \int_U \delta(x)dx = \sum_{j=1}^{\infty} (F(b_j) - F(a_j))$$

since $\delta = 0$ on C so $\int_C \delta(x)dx = 0$. Thus $m(F(U)) = m(F([a,b]))$, so that $m(F(C)) = 0$. This implies that $m(F(S)) = 0$ for any subset $S \subset C$. But since C has positive measure, it has a non-measurable subset. Then if $E = F(S)$, $m(E) = 0$ so E is measurable, but $F^{-1}(E) = S$ is not measurable.

(c)

□

Exercise 22: Suppose that F and G are absolutely continuous on $[a, b]$. Show that their product FG is also absolutely continuous. This has the following consequences.

(a) Whenever F and G are absolutely continuous in $[a, b]$,

$$\int_a^b F'(x)G(x)dx = -\int_a^b F(x)G'(x)dx + [F(x)G(x)]_a^b.$$

(b) Let F be absolutely continuous in $[-\pi, \pi]$ with $F(\pi) = F(-\pi)$. Show that if

$$a_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} F(x)e^{-inx}dx,$$

such that $F(x) \sim \sum a_n e^{inx}$, then

$$F'(x) \sim \sum ina_n e^{inx}.$$

(c) What happens if $F(-\pi) \neq F(\pi)$?

Proof. Since F and G are absolutely continuous, they are continuous and therefore bounded on the compact interval $[a, b]$. Suppose $|F|, |G| \leq M$ on this interval. Now given $\epsilon > 0$, we can choose $\delta > 0$ such that $\sum |b_j - a_j| < \delta \Rightarrow \sum |F(b_j) - F(a_j) < \frac{\epsilon}{M}$ and $\sum |G(b_j) - G(a_j)| < \frac{\epsilon}{2M}$. Then

$$\begin{aligned}
&\sum |F(b_j)G(b_j) - F(a_j)G(a_j)| \\
&\quad = \sum \frac{1}{2} |(F(b_j - F(a_j))(G(b_j) + G(a_j)) + (F(b_j) + F(a_j))(G(b_j) - G(a_j))| \\
&\quad \leq \frac{1}{2}\left(\sum |F(b_j) - F(a_j)||G(b_j) + G(a_j)| + \sum |F(b_j) + F(a_j)||G(b_j) - G(a_j)|\right) \\
&\quad \leq \frac{1}{2}\left(\sum (2M)|F(b_j) - F(a_j)| + \sum (2M)|G(b_j) - G(a_j)|\right) \\
&\quad \leq \frac{1}{2}\left(2M \cdot \frac{\epsilon}{2M} + 2M \cdot \frac{\epsilon}{2M}\right) = \epsilon.
\end{aligned}$$

This proves that FG is absolutely continuous on $[a, b]$. We now turn to the consequences of this:

(a) Since FG is absolutely continuous, it's differentiable almost everywhere. By elementary calculus, $(FG)' = F'G + FG'$ at any point where all three derivatives exist, which is almost everywhere. Integrating both sides and subtracting $\int FG'$ yields

$$\int_a^b F'(x)G(x)dx = -\int_a^b F(x)G'(x)dx + \int_a^b (FG)'(x)dx.$$

Since FG is absolutely continuous, this implies

$$\int_a^b F'(x)G(x)dx = -\int_a^b F(x)G'(x)dx + [F(x)G(x)]_a^b.$$

(b) It would be nice if the problem would actually define this for us, but I'm assuming that the $\sim$ here means "is represented by" as opposed to any kind of statement about whether the function actually converges

to its Fourier series or not. Then suppose b_n are the Fourier coefficients of F', so by definition

$$b_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} F'(x)e^{-inx}dx.$$

Using part (a), we have

$$b_n = -\frac{1}{2\pi}\int_{-\pi}^{\pi} F(x)(-ine^{-inx})dx + [F(x)e^{-inx}]_a^b = in\frac{1}{2\pi}\int_{-\pi}^{\pi} F(x)e^{-inx}dx = ina_n.$$

(c) Then all bets are off. As one example, consider $F(x) = x$ which is clearly absolutely continuous on $[-\pi, \pi]$. Then

$$a_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} xe^{-inx}dx = \frac{1}{2\pi}\left(\frac{xe^{-inx}}{-in} + \frac{e^{-inx}}{n^2}\right)\Bigg|_{-\pi}^{\pi} = \frac{2i}{n}(-1)^n$$

for $n \neq 0$, and $a_0 = \int_{-\pi}^{\pi} xdx = 0$. However, $F'(x) = 1$ which has Fourier coefficients $b_0 = 1$ and $b_n = 0$ for $n \neq 0$.

□

Exercise 25: The following shows the necessity of allowing for general exceptional sets of measure zero in the differentiation Theorems 1.4, 3.4, and 3.11. Let E be any set of measure zero in $\mathbb{R}^d$. Show that:

(a) There exists a non-negative integrable f in $\mathbb{R}^d$, such that

$$\liminf_{\substack{m(B)\to 0\\ x\in B}} \frac{1}{m(B)}\int_B f(y)dy = \infty \text{ for each } x \in E.$$

(b) When $d = 1$ this may be restated as follows. There is an increasing absolutely continuous function F such that

$$D_+F(x) = D_-F(x) = \infty, \text{ for each } x \in E.$$

Solution.

(a) Since E has measure zero, there exist open sets $\mathcal{O}_n$ with $E \subset \mathcal{O}_n$ for all n and $m(\mathcal{O}_n) < \frac{1}{2^n}$. Let $f = \sum_{n=1}^{\infty} \chi_{\mathcal{O}_n}$. Then $f \in L^1$ since

$$\int_{\mathbb{R}^d} f = \sum_{n=1}^{\infty}\int_{\mathbb{R}^d} \chi_{\mathcal{O}_n} = \sum_{n=1}^{\infty} m(\mathcal{O}_n) \leq \sum_{n=1}^{\infty}\frac{1}{2^n} = 1.$$

Now let $x \in E$. Since $\mathcal{O}_n$ is open, there exist open balls $B_n \subset \mathcal{O}_n$ with $x \in B_n$. Then for any ball $B \ni x$,

$$\int_B f(y)dy = \sum_{n=1}^{\infty} m(\mathcal{O}_n \cap B) \geq \sum_{n=1}^{\infty} m(B_n \cap B) \Rightarrow \frac{1}{m(B)}\int_B f(y)dy \geq \sum_{n=1}^{\infty}\frac{m(B_n \cap B)}{m(B)}.$$

For any N, there exists $\delta > 0$ such that $m(B) < \delta$ and $x \in B$ implies $B \subset B_j$ for all $j = 1, \ldots, N$. (This is true because $B_1 \cap \cdots \cap B_n$ is an open set containing x and hence contains an open ball around x.) Then

$$\frac{1}{m(B)}\int_B f(y)dy \geq \sum_{n=1}^{\infty}\frac{m(B_n \cap B)}{m(B)} \geq N$$

for sufficiently small B. This proves that

$$\liminf_{\substack{m(B)\to 0\\ x\in B}} \frac{1}{m(B)} \int_B f(y)dy = \infty.$$

(b) Let f be as in part (a), and

$$F(x) = \int_0^x f(y)dy.$$

Then F is absolutely continuous because it is the integral of an L^1 function; it is increasing because f is nonnegative. Now

$$D_+F(x) = \liminf_{\substack{h\to 0\\ h>0}} \frac{F(x+h)-F(x)}{h} = \liminf_{\substack{h\to 0\\ h>0}} \int_x^{x+h} f(y)dy$$

and

$$D_-F(x) = \liminf_{\substack{h\to 0\\ h<0}} \frac{F(x+h)-F(x)}{h} = \liminf_{\substack{h\to 0\\ h<0}} \int_x^{x+h} f(y)dy$$

The conclusion in (a) implies that both of these are infinite, since one can consider integrals over the balls $[x, x+h)$ and $(x-h, x]$. (Technically, I suppose we should work with open balls, but one can look at e.g. $(x-\delta, x+h)$ for sufficiently small δ.)

□

Exercise 30: A bounded function F is said to be of bounded variation on $\mathbb{R}$ if F is of bounded variation on any finite sub-interval $[a, b]$ and $\sup_{a,b} T_F(a, b) < \infty$. Prove that such an F enjoys the following two properties:

(a) $\int_{\mathbb{R}} |F(x+h) - F(x)|dx \le A|h|$, for some constant A and all $h \in \mathbb{R}$.

(b) $|\int_{\mathbb{R}} F(x)\phi'(x)dx| \le A$, where ϕ ranges over all C^1 functions of bounded support with $\sup_{x\in\mathbb{R}} |\phi(x)| \le 1$.

For the converse, and analogues in $\mathbb{R}^d$, see Problem 6* below.

Solution.

(a) First, note that it is sufficient to treat the case where F is a bounded increasing function. This is so because in general we can let $F = F_1 - F_2$ where F_1 and F_2 are bounded increasing functions; if they both satisfy the given condition, with constants A_1 and A_2, then $|F(x+h) - F(x)| \le |F_1(x+h) - F_1(x)| + |F_2(x+h) - F_2(x)|$ for all x, so $\int |F(x+h) - F(x)| \le A_1 + A_2$.
(In case someone asks why F must be a difference of bounded increasing functions, we could re-do the proof that was used on finite intervals, using the positive and negative variations of f. This avoids the problem of trying to extend from the bounded case and worrying about whether such an extension is unique.)
Suppose now that F is bounded and increasing. Since F is increasing, $|F(x+h) - F(x)| = F(x+h) - F(x)$ for $h > 0$. (By translation invariance, it is sufficient to treat the case of positive h, since $\int |F(x-h) - F(x)| = \int F(x) - F(x-h) = \int F(x+h) - F(x)$.) Then

on any interval $[a, a+h]$,

$$F(x+h)-F(x) \leq F(a+2h)-F(a) \Rightarrow \int_a^{a+h} F(x+h)-F(x) \leq h(F(a+2h)-F(a)).$$

In particular,

$$\int_{nh}^{(n+1)h} \leq h(F((n+2)h)-F(nh))$$

for any $n \in Z$. Then for any N,

$$\begin{aligned}\int_{-Nh}^{Nh} F(x+h)-F(x) &\leq h \sum_{n=-N}^{N-1} F((n+2)h)-F(nh) \\ &\leq h\Big(F((N+1)h)+F(Nh)-F((-N+1)h)-F(-Nh)\Big) \\ &\leq 2h(F(+\infty)-F(-\infty))\end{aligned}$$

since the sum telescopes. Since $F(+\infty)-F(-\infty)$ is a finite constant, this proves the result.

(b) By some algebraic legwork,

$$F(x+h)\phi(x+h)-F(x)\phi(x) = F(x)\big(\phi(x+h)-\phi(x)\big)-\phi(x+h)\big(F(x+h)-F(x)\big).$$

When we integrate both sides over $\mathbb{R}$, the left-hand side integrates to zero by translation invariance. Hence

$$\int_{\mathbb{R}} F(x)\big(\phi(x+h)-\phi(x)\big)dx = \int_{\mathbb{R}} \phi(x+h)\big(F(x+h)-F(x)\big)dx.$$

By part (a), we have

$$\begin{aligned}\Big|\int_{\mathbb{R}} F(x)\big(\phi(x+h)-\phi(x)\big)dx\Big| &= \Big|\int_{\mathbb{R}} \phi(x+h)\big(F(x+h)-F(x)\big)dx\Big| \\ &\leq \int_{\mathbb{R}} |\phi(x+h)|\big|F(x+h)-F(x)\big|dx \\ &\leq \int_{\mathbb{R}} \big|F(x+h)-F(x)\big|dx \\ &\leq A|h|.\end{aligned}$$

Hence

$$\Big|\int_{\mathbb{R}} F(x)\frac{\phi(x+h)-\phi(x)}{h}dx\Big| \leq A.$$

Now ϕ is supported on some compact set K, so ϕ' is a continuous function which is 0 outside K. Hence it has a maximum M. By the Mean Value Theorem, any difference quotient of ϕ is at most M in absolute value. Then if L is a bound for F on K, $F(x)\frac{\phi x+h-\phi(x)}{h}$ is dominated by $ML\chi_K$ for all h. Hence we can use the Dominated Convergence Theorem as $h \to 0$ to obtain

$$\Big|\int_{\mathbb{R}} F(x)\phi'(x)dx\Big| \leq A.$$

□

Exercise 31: Let F be the Cantor-Lebesgue function described in Section 3.1. Consider the curve that is the graph of F, that is, the curve given by $x(t) = t$ and $y(t) = F(t)$ with $0 \leq t \leq 1$. Prove that the length $L(\bar{x})$ of the segment $0 \leq t \leq \bar{x}$ of the curve is given by $L(\bar{x}) = \bar{x} + F(\bar{x})$. Hence the total length of the curve is 2.

Solution. It is true for any increasing function with $F(0) = 0$ that $L(\bar{x}) \leq \bar{x} + F(\bar{x})$, because for any partition $0 = t_0 < t_1 < \cdots < t_n = \bar{x}$,

$$\sum_{j=1}^{n} \sqrt{(t_j - t_{j-1})^2 + (F(t_j) - F(t_{j-1}))^2} \leq \sum_{j=1}^{n}(t_j - t_{j-1}) + (F(t_j) - F(t_{j-1})) = \bar{x} + F(\bar{x}).$$

We wish to show that this upper bound is in fact the *least* upper bound when F is the Cantor-Lebesgue function. Consider the iterates $F_n(x)$ of which this function is the limit. The interval $[0, 1]$ can be divided into $2^{n+1} - 1$ intervals on which $F_n(x)$ alternately increases and stays constant; suppose we label them $I_1, C_1, I_2, C_2, \ldots, C_{2^n-1}, I_{2^n}$. The intervals C_j have varying lengths, since they correspond to intervals that are deleted from the Cantor set at varying stages of the iteration; however, the I_j all have length $\frac{1}{3^n}$ since they correspond to the intervals remaining in the nth iteration of the Cantor set. Hence the sum of the lengths of the I_j is $\left(\frac{2^n}{3^n}\right)$, while the sum of the lengths of the C_j is $1 - \left(\frac{2^n}{3^n}\right)$.
Now let $\bar{x} \in [0, 1]$, and consider the partition P_n consisting of all points less than or equal to $\bar{x}$ which are an endpoint of one of the C_j or I_j. Thus we have $0 = t_0 < t_1 < \cdots < t_m = \bar{x}$ where F_n is increasing on $[t_0, t_1]$, constant on $[t_1, t_2]$, increasing on $[t_2, t_3]$, etc. Note also that $F(t_j) = F_n(t_j)$ since all the t_j are endpoints of the C_k intervals, which remain fixed in all successive iterations. Then

$$\begin{aligned}
&\sum_{j=1}^{m} \sqrt{(t_j - t_{j-1})^2 + (F(t_j) - F(t_{j-1}))^2} \\
=&\sum_{\substack{j=1 \\ j \text{ odd}}}^{m} \sqrt{(t_j - t_{j-1})^2 + (F(t_j) - F(t_{j-1}))^2} + \sum_{\substack{j=1 \\ j \text{ even}}}^{m} \sqrt{(t_j - t_{j-1})^2 + (F(t_j) - F(t_{j-1}))^2} \\
\geq&\sum_{\substack{j=1 \\ j \text{ odd}}}^{m} (F(t_j) - F(t_{j-1})) + \sum_{\substack{j=1 \\ j \text{ even}}}^{m} (t_j - t_{j-1}) \\
=&F(\bar{x}) + \sum_{k} |C_k \cap [0, \bar{x}]| \\
=&F(\bar{x}) + \bar{x} - \sum_{k} |I_k \cap [0, \bar{x}]| \\
\geq&F(\bar{x}) + \bar{x} - \sum_{k} |I_k| \\
=&F(\bar{x}) + \bar{x} - \left(\frac{2}{3}\right)^n.
\end{aligned}$$

Letting $n \to \infty$, this approaches $\bar{x} + F(\bar{x})$, which proves $L(\bar{x}) \geq \bar{x} + F(\bar{x})$. Since we already know $L(\bar{x}) \leq \bar{x} + F(\bar{x})$, we have $L(\bar{x}) = \bar{x} + F(\bar{x})$ as desired. □

Exercise 32: Let $f : \mathbb{R} \to \mathbb{R}$. Prove that f satisfies the Lipschitz condition

$$|f(x) - f(y)| \leq M|x - y|$$

for some M and all $x, y \in \mathbb{R}$, if and only if f satisfies the following two properties:

(i) f is absolutely continuous.
(ii) $|f'(x)| \leq M$ for a.e. x.

Solution. Suppose f is Lipschitz. Then for any $\epsilon > 0$, if we let $\delta = \frac{\epsilon}{M}$, then $\sum |b_j - a_j| < \delta \Rightarrow \sum |f(b_j) - f(a_j)| < \epsilon$. Hence f is absolutely continuous. This implies f is differentiable a.e.; if x is a point for which $f'(x)$ exists, the Lipschitz condition implies $|\frac{f(x+h)-f(x)}{h}| \leq M$ for all h. Taking the limit as $h \to 0$ implies $|f'(x)| \leq M$.

Conversely, suppose f is absolutely continuous and has bounded derivative a.e. Since absolutely continuous functions are the integrals of their derivatives,

$$|f(y) - f(x)| = \Big| \int_x^y f'(t)dt \Big| \leq \int_{\min(x,y)}^{\max(x,y)} |f'(t)|dt \leq \int_{\min(x,y)}^{\max(x,y)} M dt = M|x - y|$$

so f is Lipschitz. □

Chapter 3.6, Page 152

Problem 4: A real-valued function ϕ defined on an interval (a, b) is **convex** if the region lying above its graph $\{(x, y) \in \mathbb{R}^2 : y > \phi(x), a < x < b\}$ is a convex set. Equivalently, ϕ is convex if

$$\phi(\theta x_1 + (1 - \theta)x_2) \leq \theta\phi(x_1) + (1 - \theta)\phi(x_2)$$

for every $x_1, x_2 \in (a, b)$ and $0 \leq \theta \leq 1$. One can also observe as a consequence that we have the following inequality of the slopes:

$$\frac{\phi(x + h) - \phi(x)}{h} \leq \frac{\phi(y) - \phi(x)}{y - x} \leq \frac{\phi(y) - \phi(y - h)}{h}, \tag{1}$$

whenever $x < y$, $h > 0$, and $x + h < y$. The following can then be proved.

(a) ϕ is continuous on (a, b).
(b) ϕ satisfies a Lipschitz condition of order 1 in any proper closed subinterval $[a', b']$ of (a, b). Hence ϕ is absolutely continuous in each subinterval.
(c) ϕ' exists at all but an at most denumerable number of points, and $\phi' = D^+\phi$ is an increasing function with

$$\phi(y) - \phi(x) = \int_x^y \phi'(t)dt.$$

(d) Conversely, if ψ is any increasing function on (a, b), then $\phi(x) = \int_c^x \psi(t)dt$ is a convex function in (a, b) for $c \in (a, b)$.

Solution.

(a) Suppose to the contrary that ϕ is discontinuous at some $x \in (a, b)$. This means there exists $\epsilon > 0$ and an infinite sequence $x_n \to x$ with $x_n \in (a, b)$ and $|\phi(x_n) - \phi(x)| > \epsilon$ for all n. Since the sequence x_n is infinite, it must have infinitely many points in one of the following categories:
  (i) $\phi(x_n) > \phi(x) + \epsilon$
  (ii) $\phi(x_n) < \phi(x) - \epsilon$

We will treat each case separately and obtain a contradiction.

  (i) Assume WLOG that the entire sequence x_n is in this category (otherwise, take a subsequence that is). We may also assume x_n converges to x monotonically, since otherwise we can again take a subsequence. Let $L(\theta) = \theta\phi(x) + (1 - \theta)\phi(x_1)$ for $0 \le \theta \le 1$. This is a continuous function of θ, so $\exists\delta > 0$ such that $|\theta| < \delta \Rightarrow L(\theta) < \phi(x) + \epsilon$. Now let $\theta_n = \frac{x_n - x_1}{x - x_1}$. Then $0 \le \theta_n \le 1$ since $x_n \to x$ monotonically. Note also that $x_n = \theta_n x + (1 - \theta_n)x_1$. Now $\theta_n \to 0$ as $n \to \infty$, so $\theta_n < \delta$ for n sufficiently large. But this implies

$$\phi(x_n) = \phi(\theta_n x + (1 - \theta_n)x_1) \le L(\theta_n) < \phi(x) + \epsilon,$$

  for sufficiently large n, a contradiction.

  (ii) Again, we assume WLOG that $x_n \to x$ monotonically. Let $y \in (a, b)$ such that x is between y and x_n for all n. Then $\theta_n = \frac{x_n - x}{x_n - y}$ has the properties that $0 \le \theta_n \le 1$ and $\theta_n \to 1$ as $n \to \infty$. Now $x = \theta_n x_n + (1 - \theta_n)y$, so

$$\phi(x) \le \theta_n\phi(x_n) + (1 - \theta_n)\phi(y)$$

  for all n. But $\theta_n \to 1$, and since $\phi(x_n) < \phi(x) - \epsilon$, this implies $\phi(x) < \phi(x) - \epsilon$ for sufficiently large n, a contradiction.

Hence ϕ is continuous.

(b) First, I prove an inequality of slopes that I like better than the one given. I claim that for $s < t < u$ with $s, t, u \in (a, b)$,

$$\frac{\phi(t) - \phi(s)}{t - s} \le \frac{\phi(u) - \phi(s)}{u - s} \le \frac{\phi(u) - \phi(t)}{u - t}. \tag{2}$$

This follows straightforwardly from the convexity condition:

$$\begin{aligned}
&t = \frac{u - t}{u - s}s + \frac{t - s}{u - s}u \\
&\Rightarrow \phi(t) \le \frac{u - t}{u - s}\phi(s) + \frac{t - s}{u - s}\phi(u) \\
&\Rightarrow (u - s)\phi(t) \le (u - t)\phi(s) + (t - s)\phi(u) \qquad (3)\\
&\Rightarrow (u - s)\phi(t) - (u - s)\phi(s) \le (s - t)\phi(s) + (t - s)\phi(u) \\
&\Rightarrow \frac{\phi(t) - \phi(s)}{t - s} \le \frac{\phi(u) - \phi(s)}{u - s}.
\end{aligned}$$

Taking a different route from inequality (3) leads to

$$\begin{aligned}
(u - t)\phi(u) - (u - t)\phi(s) &\le (u - s)\phi(u) - (u - s)\phi(t) \\
\Rightarrow \frac{\phi(u) - \phi(s)}{u - s} &\le \frac{\phi(u) - \phi(t)}{u - t}
\end{aligned}$$

as desired. Given this inequality of slopes, we can easily prove that ϕ is Lipschitz on $[a', b']$. Choose $h > 0$ such that $[a' - h, b' + h] \subset (a, b)$. Then for $x, y \in [a', b']$, suppose WLOG that $x < y$; since $a' - h < a' \leq x < y \leq b' < b' + h$, the slope inequality yields

$$\frac{\phi(a'-h) - \phi(a')}{h} \leq \frac{\phi(y) - \phi(a'-h)}{y - a' + h} \leq \frac{\phi(y) - \phi(x)}{y - x} \leq \frac{\phi(b'+h) - \phi(x)}{b' + h - x} \leq \frac{\phi(b'+h) - \phi(b')}{h}.$$

The leftmost and rightmost terms above are constants, which we may call m and M; we thus have $m|y - x| \leq |\phi(y) - \phi(x)| \leq M|y - x|$, whence ϕ is Lipschitz on $[a', b']$.

(c) Since ϕ is Lipschitz on any closed subinterval $[x, y] \subset (a, b)$, ϕ is absolutely continuous on $[x, y]$ by Exercise 32 above; hence $\phi(y) - \phi(x) = \int_x^y \phi'(t)dt$.
Now inequality (2) implies that $\frac{\phi(x+h)-\phi(x)}{h}$ is an increasing function of h at any $x \in (a, b)$. This implies that $D^+ = D_+$ and $D^- = D_-$, that $|D^+|, |D^-| < \infty$, and that $D^+ \geq D^-$. The inequality (1) tells us that $x < y \Rightarrow D^+\phi(x) \leq D^-\phi(y)$. This in turn implies that D^+ and D^- are increasing. To show that $D^+ = D^-$ except at countably many points, let $\{x_\alpha\}$ be those points in (a, b) for which this is not true, and define $j_\alpha > 0$ by $j_\alpha = D^+(\phi)(x_\alpha) - D^-(\phi)(x_\alpha)$. Then on any subinterval $[a', b']$, if $\{x_1, \ldots, x_n\} \subset [a', b']$, we have

$$\begin{aligned}
&\sum_{k=1}^{n+1} \Big(D^+(\phi)(x_k) - D^-(\phi)(x_k)\Big) \\
\leq &\sum_{k=1}^{n+1} \Big(D^+(\phi)(x_k) - D^-(\phi)(x_k)\Big) + \sum_{k=1}^{n+1} \Big(D^-(\phi)(x_k) - D^+(\phi)(x_{k-1})\Big) \\
= &D^+(\phi)(b') - D^-(\phi)(a'),
\end{aligned}$$

where we use the convention $x_0 = a'$ and $x^{n+1} = b'$. This implies that

$$\sum_{x \in [a', b']} \Big(D^+(\phi)(x_k) - D^-(\phi)(x_k)\Big)$$

is finite, because all finite sub-sums are bounded by the finite constant $D^+(\phi)(b') - D^-(\phi)(a')$. So this sum can containly only countably many nonzero terms, which means only countably many points in $[a', b']$ can have $D^-(\phi)(x) \neq D^+(\phi)(x)$. Since (a, b) is a countable union of closed subintervals (e.g. $\cap[a + \frac{1}{n}, b - \frac{1}{n}]$), it can contain only countably many points for which $D^+ \neq D^-$. Everywhere else, the derivative exists.

(d) Because ψ is increasing,

$$\begin{aligned}
\phi\big(\theta x_1 + (1-\theta)x_2\big) &= \int_c^{\theta x_1+(1-\theta)x_2} \psi(t)dt \\
&= \int_c^{x_2} \psi(t)dt - \int_{\theta x_1+(1-\theta)x_2}^{x_2} \psi(t)dt \\
&= \int_c^{x_2} \psi(t)dt - \theta \int_{\theta x_1+(1-\theta)x_2}^{x_2} \psi(t)dt - (1-\theta)\int_{\theta x_1+(1-\theta)x_2}^{x_2} \psi(t)dt \\
&\leq \int_c^{x_2} \psi(t)dt - \theta \int_{\theta x_1+(1-\theta)x_2}^{x_2} \psi(t)dt - (1-\theta)\theta(x_2-x_1)\psi(\theta x_1+(1-\theta)x_2) \\
&\leq \int_c^{x_2} \psi(t)dt - \theta \int_{\theta x_1+(1-\theta)x_2}^{x_2} \psi(t)dt - \theta\int_{x_1}^{\theta x_1+(1-\theta)x_2} \psi(t)dt \\
&= \int_c^{x_2} \psi(t)dt - \theta \int_{x_1}^{x_2} \psi(t)dt \\
&= \theta \int_c^{x_1} \psi(t)dt + (1-\theta)\int_c^{x_2} \psi(t)dt \\
&= \theta\phi(x_1) + (1-\theta)\phi(x_2).
\end{aligned}$$

So ϕ is convex.

□

Chapter 4.7, Page 193

Exercise 4: Prove from the definition that $\ell^2(\mathbb{Z})$ is complete and separable.

Solution. The proof that ℓ^2 is complete is exactly the same as the proof that L^2 is complete, from pp. 159-160 of the textbook. Let $\{a_j^{(m)}\}_{m=1}^{\infty}$ be a Cauchy sequence in $\ell^2(\mathbb{Z})$. For each $k \geq 1$ we can choose n_k such that $m, n \geq n_k \Rightarrow \|a^{(m)} - a^{(n)}\| < \frac{1}{2^k}$ and $n_k < n_{k+1}$. Then the subsequence $a^{(n_k)}$ has the property that $\|a^{(n_{k+1})} - a^{(n_k)}\| \leq \frac{1}{2^k}$. Define sequences $a = \{a_j\}$ and $b = \{b_j\}$ by

$$a_j = a_j^{(n_1)} + \sum_{k=1}^{\infty} \left(a_j^{(n_{k+1})} - a_j^{(n_k)}\right)$$

and

$$b_j = |a_j^{(n_1)}| + \sum_{k=1}^{\infty} \left|a_j^{(n_{k+1})} - a_j^{(n_k)}\right|$$

and the partial sums

$$S_j^{(a,K)} = a_j^{(n_1)} + \sum_{k=1}^{K} \left(a_j^{(n_{k+1})} - a_j^{(n_k)}\right)$$

and

$$S_j^{(b,K)} = |a_j^{(n_1)}| + \sum_{k=1}^{K} \left|a_j^{(n_{k+1})} - a_j^{(n_k)}\right|.$$

Then
$$\|S^{(b,K)}\| \leq \|a^{(n_1)}\| + \sum_{k=1}^{K} \frac{1}{2^k}$$
by the triangle inequality; letting $K \to \infty$, $\|b\|$ converges by the monotone convergence theorem (for sums, but hey, sums are just integrals with discrete measures), so $\|a\|$ converges since it converges absolutely. Of course, this implies that b_j and hence a_j converges for each j; since the partial sums are $S^{(a,K)} = a^{(n_{k+1})}$ by construction, $a_j^{(n_{k+1})} \to a_j$ for all j. Now given $\epsilon > 0$, choose N such that $\|a^{(n)} - a^{(m)}\| < \frac{\epsilon}{2}$ for $n, m > N$, and let $n_K > N$ such that $\|a^{(n_K)} - a\| < \frac{\epsilon}{2}$. Then
$$m > N \Rightarrow \|a^{(m)} - a\| \leq \|a^{(m)} - a^{(n_K)}\| + \|a^{(n_K)} - a\| < \epsilon.$$
Hence $a^{(m)} \to a$.

To prove that ℓ^2 is separable, consider the subset $\mathcal{D}$ consisting of all rational sequences which are 0 except at finitely many values. This is countable because
$$\begin{aligned}\mathcal{D} &= \{\text{rational sequences of finite length}\}\\ &= \bigcup_{N=1}^{\infty} \{\text{sequences } \{a_n\} \text{ with } a_n \in \mathbb{Q} \text{ and } a_n = 0 \text{ for } |n| > N\}\\ &= \bigcup_{N=1}^{\infty} \mathbb{Q}^{2N+1}\end{aligned}$$
is a countable union of countable sets. Now let $\{b_n\} \in \ell^2$ be any square summable sequence. Given $\epsilon > 0$, there exists N such that
$$\sum_{|n|>N} |b_n|^2 < \frac{\epsilon^2}{2}$$
since the infinite sum converges. Then for each $j = -N, \ldots, N$, we can choose a rational number q_j with $|q_j - b_j|^2 < \frac{\epsilon^2}{2^{2+N+j}}$. If we also define $q_j = 0$ for $|j| > N$, then $q \in \mathcal{D}$ and
$$\|q - b\|^2 = \sum_{j=-\infty}^{\infty} |q_j - b_j|^2 = \sum_{|j|>N} |b_j - 0|^2 + \sum_{j=-N}^{N} |q_j - b_j|^2 < \frac{\epsilon^2}{2} + \sum_{s=0}^{2N} \frac{\epsilon^2}{2^{2+s}} < \epsilon^2.$$
This shows that $\mathcal{D}$ is dense. □

Exercise 5: Establish the following relations between $L^2(\mathbb{R}^d)$ and $L^1(\mathbb{R}^d)$:

(a) Neither the inclusion $L^2(\mathbb{R}^d) \subset L^1(\mathbb{R}^d)$ nor the inclusion $L^1(\mathbb{R}^d) \subset L^2(\mathbb{R}^d)$ is valid.

(b) Note, however, that if f is supported on a set E of finite measure and if $f \in L^2(\mathbb{R}^d)$, applying the Cauchy-Schwarz inequality to $f\chi_E$ gives $f \in L^1(\mathbb{R}^d)$, and
$$\|f\|_1 \leq m(E)^{1/2} \|f\|_2.$$

(c) If f is bounded ($|f(x)| \leq M$), and $f \in L^1(\mathbb{R}^d)$, then $f \in L^2(\mathbb{R}^d)$ with
$$\|f\|_2 \leq M^{1/2} \|f\|_1^{1/2}.$$

Solution.

(a) Let $f(x) = \chi_{|x|\geq 1}\frac{1}{|x|^{d/2}}$ and $g(x) = \chi_{|x|\leq 1}\frac{1}{|x|^d}$. Then Exercise 10 of Chapter 2 shows that f and g^2 are integrable, but f^2 and g are not.

(b) Applying the Cauchy-Schwarz inequality to the inner product of $f\chi_E$ and χ_E,

$$\|f\|_1 = \|f\chi_E\|_1 = \int |f\chi_E| \leq \left(\int |f|\right)^{1/2}\left(\int \chi_E\right)^{1/2} = m(E)^{1/2}\|f\|_2.$$

(c) Since $|f| \leq M$,

$$|f|^2 \leq M|f| \Rightarrow \|f\|_2^2 = \int |f|^2 \leq M\int |f| = M\|f\|_1 \Rightarrow \|f\|_2 \leq M^{1/2}\|f\|_1^{1/2}.$$

□

Exercise 6: Prove that the following are dense subspaces of $L^2(\mathbb{R}^d)$:

(a) The simple functions.

(b) The continuous functions of compact support.

Solution.

(a) It is sufficient to treat the case of nonnegative f, since every complex L^2 function is a linear combination of nonnegative L^2 functions. We know there exists a sequence of simple functions $s_n \nearrow f$ with $0 \leq s_n \leq f$. Then $|f - s_n|^p \leq |f|^p$ so by the Dominated Convergence Theorem, $\int |f - s_n|^p \to 0$. Hence $s_n \to f$ in L^p. Therefore the simple functions are dense.

(b) Let $s \in L^p(\mathbb{R}^d)$ be a simple function. It is sufficient to find $g \in C_C(\mathbb{R}^d)$ with $\|g - s\|_p < \epsilon$. If $s = 0$, $s \in C_C(\mathbb{R}^d)$ and we're done. Otherwise, since s is simple, $0 < \|s\|_\infty < \infty$; since it is in L^p, it must be supported on a set E of finite measure. Now Lusin's theorem enables us to construct $g \in C_C(\mathbb{R}^d)$ with $m(\{g \neq s\}) < \left(\frac{\epsilon}{2\|s\|_\infty}\right)^p$ and $\sup|g| \leq \sup|s| < \infty$. (To do this, construct g from Lusin's theorem; then if $g \geq \|s\|_\infty$ on the closed set F, change g to $\|s\|_\infty$ on F. See Rudin, *Real and Complex Analysis* pp. 55-56.) Then $|g - s| \leq 2\|s\|_\infty$ and is nonzero on a set of measure $\left(\frac{\epsilon}{2\|s\|_\infty}\right)^p$, so

$$\int |g - s|^p \leq 2^p\|s\|_\infty\left(\frac{\epsilon}{2\|s\|_\infty}\right)^p = \epsilon^p \Rightarrow \|g - s\|_p < \epsilon.$$

□

Exercise 7: Suppose $\{\phi_k\}_{k=1}^\infty$ is an orthonormal basis for $L^2(\mathbb{R}^d)$. Prove that the collection $\{\phi_{k,j}\}_{1\leq k,j<\infty}$ with $\phi_{k,j}(x,y) = \phi_k(x)\phi_j(y)$ is an orthonormal basis of $L^2(\mathbb{R}^d \times \mathbb{R}^d)$.

Solution. First, note that $\phi_{k,j}$ is indeed in $L^2(\mathbb{R}^d \times \mathbb{R}^d$, since

$$\int_{\mathbb{R}^{2d}} |\phi(k,j)|^2 = \int_{\mathbb{R}^{2d}} |\phi_k(x)|^2|\phi_j(y)|^2 dxdy = \left(\int_{\mathbb{R}^d} |\phi_k(x)|^2 dx\right)\left(\int_{\mathbb{R}^d} |\phi_j(y)|^2 dy\right) = 1$$

by Fubini's Theorem. Also,

$$\begin{aligned}
\langle \phi_{j,k}\phi_{\ell,m}\rangle &= \int_{\mathbb{R}^{2d}} \phi_{j,k}\overline{\phi_{\ell,m}} \\
&= \int_{\mathbb{R}^{2d}} \phi_j(x)\phi_k(y)\overline{\phi_\ell(x)\phi_m(y)}dxdy \\
&= \left(\int_{\mathbb{R}^d} \phi_j(x)\overline{\phi_\ell(x)}dx\right)\left(\int_{\mathbb{R}^d} \phi_k(y)\overline{\phi_m(y)}dy\right) \\
&= \delta_j^\ell \delta_k^m
\end{aligned}$$

so $\{\phi_{j,k}\}$ is an orthonormal set. To show that linear combinations of $\{\phi_{j,k}\}$ are dense in $L^2(\mathbb{R}^d)$, one approach would be to blow this problem out of the water with the Stone-Weierstrass theorem. We know $C_C(\mathbb{R}^d)$ is dense in $L^2(\mathbb{R}^d)$, and the Stone-Weierstrass theorem tells us that linear combinations of "separable" continuous functions (i.e. functions of the form $\sum_{i=1}^m f_i(x)g_i(y)$) are dense in $C_C(\mathbb{R}^d)$. It is easy to verify that functions of this form can be approximated by $\{\phi_{j,k}\}$, so we're done.
Alternatively, we can follow the approach given in the hint. Let $f \in L^2(\mathbb{R}^{2d})$ and suppose that $\langle f, \phi_{j,k}\rangle = 0$ for all j and k. By Fubini's Theorem,

$$\begin{aligned}
0 &= \int_{\mathbb{R}^{2d}} f(x,y)\overline{\phi_{j,k}(x,y)} \\
&= \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} f(x,y)\overline{\phi_{j,k}(x,y)}dy\right)dx \\
&= \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} f(x,y)\overline{\phi_k(y)}dy\right)\overline{\phi_j(x)}dx
\end{aligned}$$

Hence, if we define

$$f_k(x) = \int_{\mathbb{R}^d} f(x,y)\overline{\phi_k(y)}dy,$$

we see that

$$\int_{\mathbb{R}^d} f_k(x)\overline{\phi_j(x)} = 0$$

for all j. Because $\{\phi_j\}$ is an orthonormal basis, this implies that $f_k(x) = 0$ for all k. Because $\{\phi_k\}$ is an orthonormal basis, this in turn implies that $f(x,y) = 0$. Since $f \perp \phi_{j,k} \Rightarrow f = 0$, $\{\phi_{j,k}\}$ is an orthonormal basis. □

Exercise 9: Let $\mathcal{H}_1 = L^2([-\pi,\pi])$ be the Hilbert space of functions $F(e^{i\theta})$ on the unit circle with inner product $(F,G) = \frac{1}{2\pi}\int_{-\pi}^{\pi} F(e^{i\theta})\overline{G(e^{i\theta})}d\theta$. Let $\mathcal{H}_2$ be the space $L^2(\mathbb{R})$. Using the mapping

$$x \mapsto \frac{i-x}{i+x}$$

of $\mathbb{R}$ to the unit circle, show that:

(a) The correspondence $U : F \to f$, with

$$f(x) = \frac{1}{\pi^{1/2}(i+x)}F\left(\frac{i-x}{i+x}\right)$$

gives a unitary mapping of $\mathcal{H}_1$ to $\mathcal{H}_2$.

(b) As a result,

$$\left\{ \frac{1}{\pi^{1/2}} \left(\frac{i-x}{i+x} \right)^n \frac{1}{i+x} \right\}_{n=-\infty}^{\infty}$$

is an orthonormal basis of $L^2(\mathbb{R})$.

Solution.

(a) If we define $\theta = 2\tan^{-1}(x)$, then $x = \tan\left(\frac{\theta}{2}\right)$, $\frac{i-x}{i+x} = e^{i\theta}$, $1 + x^2 = \sec^2\left(\frac{\theta}{2}\right)$, and $dx = \frac{1}{2}\sec^2\left(\frac{\theta}{2}\right) d\theta$. (Brings back memories of high school calculus, don't it?) Then

$$\begin{aligned}
\int_{\mathbb{R}} |f(x)|^2 dx &= \int_{\mathbb{R}} \frac{1}{\pi|i+x|^2} \left| F\left(\frac{i-x}{i+x} \right) \right|^2 dx \\
&= \int_{\mathbb{R}} \frac{1}{\pi} \frac{1}{x^2+1} \left| F\left(\frac{i-x}{i+x} \right) \right|^2 dx \\
&= \int_{-\pi}^{\pi} \frac{1}{\pi} \frac{1}{\sec^2(\theta/2)} |F(e^{i\theta})|^2 \frac{1}{2} \sec^2\left(\frac{\theta}{2}\right) d\theta \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} |F(e^{i\theta})|^2 d\theta
\end{aligned}$$

so $\|f\|_{\mathcal{H}_2} = \|F\|_{\mathcal{H}_1}$. So U is unitary.

(b) By the Riesz-Fischer theorem, $\{e^{in\theta}\}$ is an orthonormal basis for $L^2(T)$. Because U is unitary,

$$\{U(e^{in\theta})\} = \left\{ \frac{1}{\sqrt{\pi}} \left(\frac{i-x}{i+x} \right)^n \frac{1}{i+x} \right\}$$

is an orthonormal basis for $L^2(\mathbb{R})$.

□

Exercise 10: Let $\mathcal{S}$ denote a subspace of a Hilbert space $\mathcal{H}$. Prove that $(\mathcal{S}^\perp)^\perp$ is the smallest closed subspace of $\mathcal{H}$ that contains $\mathcal{S}$.

Solution. Let

$$\bar{S} = \bigcap_{\substack{V \subset \mathcal{H} \text{ subspace} \\ V \text{ closed}}} V.$$

Then $\bar{S}$ is a closed subspace, because the intersection of closed sets is closed and the intersection of subspaces is a subspace. It is obviously the smallest closed subspace containing S. We want to show that $\bar{S} = (S^\perp)^\perp$. Clearly $\bar{S} \subset (S^\perp)^\perp$ since the latter is a closed subspace containing S. To establish the reverse inclusion, we first show that $(\bar{S})^\perp = S^\perp$. Clearly $(\bar{S})^\perp \subset S^\perp$ since $S \subset \bar{S}$. To show the opposite inclusion, let $x \in S^\perp$. Then for any $w \in \bar{S}$, there is a sequence $w_n \subset S$ with $w_n \to w$. Because inner products are continuous, $0 = \langle w_n, x \rangle \to \langle w, x \rangle$ so $x \perp w$ for any $w \in \bar{S}$. This proves $S^\perp \subset (\bar{S})^\perp$. Since they're equal, we can use Proposition 4.2 to write

$$\mathcal{H} = \bar{S} \oplus S^\perp.$$

Now let $x \in (S^\perp)^\perp$. Then we can write $x = v + w$ where $v \in \bar{S}$ and $w \in S^\perp$. Then

$$\langle x, w \rangle = \langle v, w \rangle + \langle w, w \rangle = \langle w, w \rangle.$$

But $\langle x, w \rangle = 0$ because $w \in S^\perp$ and $x \in (S^\perp)^\perp$. Hence $w = 0$ and $x \in \bar{S}$. $\square$

Exercise 11: Let P be the orthogonal projection associated with a closed subspace $\mathcal{S}$ in a Hilbert space $\mathcal{H}$, that is,

$$P(f) = f \text{ if } f \in \mathcal{S} \text{ and } P(f) = 0 \text{ if } f \in \mathcal{S}^\perp.$$

(a) Show that $P^2 = P$ and $P^* = P$.
(b) Conversely, if P is any bounded operator satisfying $P^2 = P$ and $P^* = P$, prove that P is the orthogonal projection for some closed subspace of $\mathcal{H}$.
(c) Using P, prove that if $\mathcal{S}$ is a closed subspace of a separable Hilbert space, then $\mathcal{S}$ is also a separable Hilbert space.

Solution.
(a) Let $x \in \mathcal{H}$ and write $x = x_S + x_{S^\perp}$. Then

$$P^2(x) = P(P(x)) = P(x_S) = x_S = P(x)$$

so $P^2 = P$. Moreover, if $y = y_S + y_{S^\perp}$ is any other vector in $\mathcal{H}$, then

$$\langle Px, y \rangle = \langle x_S, y_S + y_{S^\perp} \rangle = \langle x_S, y_S \rangle = \langle x_S + x_{S^\perp}, y_S \rangle = \langle x, Py \rangle$$

so $P = P^*$.
(b) Let $S = \text{im}(P)$ which is a subspace of $\mathcal{H}$. To show S is closed, suppose $x_n \in S$ and $x_n \to x$. Then because P is bounded, it's continuous, so $Px_n \to Px$. But $Px_n = x_n$, so $x_n \to Px$ which implies $Px = x$. Hence $x \in S$, so S is closed. Also, if $w \in S^\perp$, then for all $v \in S$,

$$\langle x, Pw \rangle = \langle Px, w \rangle = \langle x, w \rangle = 0$$

so $Pw \in S^\perp$. But $Pw \in S$, so $Pw = 0$. Now using Proposition 4.2, if y is any vector in $\mathcal{H}$, then

$$y = y_S + y_{S^\perp}.$$

Then by linearity, $P(y) = P(y_S) + P(y_{S^\perp}) = y_S + 0 = y_S$. So P does the same thing to y as orthogonal projection onto S, for any $y \in \mathcal{H}$.
(c) Let $\mathcal{H}$ be a separable Hilbert space, $\{\phi_n\}$ a countable dense set, and S a closed subspace. I claim that $\{P_S\phi_n\}$ is dense in S. For any $x \in S$, we can find a sequence $\phi_k \to x$. Then $P_S(\phi_k - x) = P_S\phi_k - x$ since $Px = x$. Since projections do not increase length,

$$\|P_S\phi_k - x\| \leq \|\phi_k - x\| \Rightarrow (P_S\phi_k) \to x.$$

$\square$

Exercise 12: Let E be a measurable subset of $\mathbb{R}^d$, and suppose $\mathcal{S}$ is the subspace of $L^2(\mathbb{R}^d)$ of functions that vanish for a.e. $x \notin E$. Show that the orthogonal projection P on $\mathcal{S}$ is given by $P(f) = \chi_E \cdot f$, where χ_E is the characteristic function of E.

Solution. Define a linear operator $T : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ by $T(f) = \chi_E f$. Then $T^2(f) = \chi_E^2 f = \chi_E f = T(f)$. Moreover,

$$\langle Tf, g \rangle = \int_{\mathbb{R}^d} \chi_E f \bar{g} = \int_{\mathbb{R}^d} f \overline{\chi_E g} = \langle f, Tg \rangle$$

so $T^* = T$. We also note that T is bounded since $|Tf(x)| \leq |f(x)|$ for all x, so $\|Tf\| \leq \|f\|$. By problem 11c, T is a projection onto its image. But $\operatorname{im}(T)$ is precisely those functions which are 0 a.e. on E^c. Hence T is the desired projection. □

Exercise 13: Suppose P_1 and P_2 are a pair of orthogonal projections on S_1 and S_2, respectively. Then P_1P_2 is an orthogonal projection if and only if P_1 and P_2 commute, that is, $P_1P_2 = P_2P_1$. In this case, P_1P_2 projects onto $S_1 \cap S_2$.

Solution. Suppose P_1P_2 is an orthogonal projection. Then

$$P_2P_1 = P_2^*P_1^* = (P_1P_2)^* = P_1P_2$$

so they commute. On the other hand, suppose they commute; then

$$(P_1P_2)^2 = P_1^2P_2^2 = P_1P_2$$

and

$$(P_1P_2)^* = P_2^*P_1^* = P_2P_1 = P_1P_2$$

so P_1P_2 is an orthogonal projection. The image of P_1P_2 is a subspace of S_1 because $P_1P_2v = P_1(P_2v) \in S_1$ for any v; similarly, it's a subspace of S_2 because $P_1P_2v = P_2(P_1v) \in S_2$. Hence the image of P_1P_2 is a subspace of $S_1 \cap S_2$. But every vector in $S_1 \cap S_2$ is fixed by both P_1 and P_2, and hence by P_1P_2. Thus, the image of P_1P_2 is precisely $S_1 \cap S_2$. □

Exercise 14: Suppose $\mathcal{H}$ and $\mathcal{H}'$ are two completions of a pre-Hilbert space $\mathcal{H}_0$. Show that there is a unitary mapping from $\mathcal{H}$ to $\mathcal{H}'$ that is the identity on $\mathcal{H}_0$.

Solution. Define $U : \mathcal{H} \to \mathcal{H}'$ as follows: Given $x \in \mathcal{H}$, choose a sequence $f_n \to x$ with $f_n \in \mathcal{H}_0$. Define $U(x) = \lim f_n$ in $\mathcal{H}'$. (This limit exists because f_n is Cauchy.) To show this is well-defined, suppose $g_n \in \mathcal{H}_0$ is another sequence converging to x. Then $f_n - g_n \to 0$ in both spaces (because $f_n - g_n \in \mathcal{H}_0$ for all n), so they have the same limit in $\mathcal{H}'$ as well. This shows that U is well-defined. Clearly it is the identity on $\mathcal{H}_0$. To show that it's unitary, we need only use the continuity of the norm: Suppose $f_n \to x$ in $\mathcal{H}$ and $f_n \to U(x)$ in $\mathcal{H}'$. Then $\|x\| = \lim \|f_n\| = \|U(x)\|$. Hence U is unitary. □

Exercise 15: Let T be any linear transformation from $\mathcal{H}_1$ to $\mathcal{H}_2$. If we suppose that $\mathcal{H}_1$ is finite-dimensional, then T is automatically bounded.

Solution. Let $\{e_1, \ldots, e_n\}$ be an orthonormal basis for $\mathcal{H}_1$. (This is a basis in the usual linear algebra sense, i.e. every vector is a *finite* linear combination of the basis vectors.) Let $m_i = \|T(e_i)\|$. Then if $x \in \mathcal{H}_1$ is a unit vector, we can write $x = \sum_{i=1}^n c_i e_i$ with $\sum_{i=1}^n |c_i|^2 = 1$. By the Triangle Inequality,

$$\|T(x)\| = \left\| T\left(\sum_{i=1}^n c_i e_i\right) \right\| = \left\| \sum_{i=1}^n c_i T(e_i) \right\| \leq \sum_{i=1}^n |c_i| \|T(e_i)\| \leq nM$$

where $M = \max(m_i)$. (In fact, this bound can be improved to $M\sqrt{n}$, because $\sum |c_i| \leq \sqrt{n}$ by the Cauchy-Schwarz inequality.) □

Exercise 18: Let $\mathcal{H}$ denote a Hilbert space, and $\mathcal{L}(\mathcal{H})$ the vector space of all bounded linear operators on $\mathcal{H}$. Given $T \in \mathcal{L}(\mathcal{H})$, we define the operator norm

$$\|T\| = \inf\{B : \|Tv\| \leq B\|v\|, \text{ for all } v \in \mathcal{H}\}.$$

(a) Show that $\|T_1 + T_2\| \leq \|T_1\| + \|T_2\|$ whenever $T_1, T_2 \in \mathcal{L}(\mathcal{H})$.
(b) Prove that

$$d(T_1, T_2) = \|T_1 - T_2\|$$

defines a metric on $\mathcal{L}(\mathcal{H})$.
(c) Show that $\mathcal{L}(\mathcal{H})$ is complete in the metric d.

Solution.
(a) This is easier if we use another expression for $\|T\|$, such as

$$\|T\| = \sup_{xin\mathcal{H}, x\neq 0} \frac{\|Tx\|}{\|x\|}.$$

If we call this supremum S, then we have $\|Tx\| \leq S\|x\|$ for all x; moreover, for any $\alpha < S$, there exists x with $\|Tx\|/\|x\| > \alpha \Rightarrow \|Tx\| > \alpha\|x\|$. Hence S is the infimum (in fact, the minimum) of such bounds, so this definition really is equivalent. Then

$$\|T_1 + T_2\| = \sup \frac{\|T_1 x\|}{\|x\|} + \frac{\|T_2 x\|}{\|x\|} \leq \sup \frac{\|T_1 x\|}{\|x\|} + \sup \frac{\|T_2 x\|}{\|x\|} = \|T_1\| + \|T_2\|.$$

(b) Since $\|Tx\| = \|-Tx\|$ for all x, $\|T\| = \|-T\|$. Then

$$d(T_1, T_2) = \|T_1 - T_2\| = \|T_2 - T_1\| = d(T_2, T_1).$$

Clearly $d(T_1, T_1) = 0$; conversely, if $T_1 \neq T_2$ then there exists some x with $\|(T_1 - T_2)x\| > 0$, so $\|T_1 - T_2\| > 0$. Finally, using part (a),

$$d(T_1, T_3) = \|T_1 - T_3\| = \|(T_1 - T_2) + (T_2 - T_3)\| \leq \|T_1 - T_2\| + \|T_2 - T_3\| = d(T_1, T_2) + d(T_2, T_3).$$

Hence d is a metric.
(c) Let $\{T_n\}$ be a Cauchy sequence in $\mathcal{L}(\mathcal{H})$. Define $T(x) = \lim T_n(x)$ for all $x \in \mathcal{H}$. This limit exists because $\|(T_m - T_n)x\| \leq \|T_m - T_n\|\|x\|$ so $\{T_n x\}$ is a Cauchy sequence in $\mathcal{H}$. T is linear by the linearity of limits. Finally, since $T_n x \to Tx$, the continuity of the norm implies $\|T_n x\| \to \|Tx\|$ for all x. Hence $\|T\| = \lim \|T_n\|$ which is finite because $|(\|T_m\| - \|T_n\|)| \leq \|T_m - T_n\|$ by the triangle inequality, so $\|T_n\|$ is a Cauchy sequence of real numbers. So T is bounded.

□

Exercise 19: If T is a bounded linear operator on a Hilbert space, prove that

$$\|TT^*\| = \|T^*T\| = \|T\|^2 = \|T^*\|^2.$$

Solution. We already know $\|T\|^2 = \|T^*\|^2$ from Proposition 5.4. Now

$$\begin{aligned}\|T^*T\| &= \sup_{\|f\|=\|g\|=1} |\langle T^*Tf, g\rangle| \\ &= \sup_{\|f\|=\|g\|=1} |\langle Tf, Tg\rangle| \\ &\leq \sup_{\|f\|=\|g\|=1} \|Tf\|\|Tg\| \\ &= \sup_{\|f\|=1} \|Tf\| \sup_{\|g\|=1} \|Tg\| \\ &= \|T\|^2.\end{aligned}$$

To show that equality is achieved, choose a sequence f_n with $\|f_n\| = 1$ and $\|Tf_n\| \to \|T\|$. Then $\langle Tf_n, Tf_n\rangle \to \|T\|^2$, so

$$\|T^*T\| = \sup_{\|f\|=\|g\|=1} \langle Tf, Tg\rangle \geq \sup_{\|f\|=1} \langle Tf, Tf\rangle \geq \|T\|^2.$$

Hence $\|T^*T\| = \|T^2\|$. Finally, replacing T with T^* yields $\|TT^*\| = \|T^*\|^2 = \|T\|^2$ and we are done. □

Exercise 20: Suppose $\mathcal{H}$ is an infinite-dimensional Hilbert space. We have seen an example of a sequence $\{f_n\}$ in $\mathcal{H}$ with $\|f_n\| = 1$ for all n, but for which no subsequence of $\{f_n\}$ converges in $\mathcal{H}$. However, show that for any sequence $\{f_n\}$ in $\mathcal{H}$ with $\|f_n\| = 1$ for all n, there exist $f \in \mathcal{H}$ and a subsequence $\{f_{n_k}\}$ such that for all $g \in \mathcal{H}$, one has

$$\lim_{n\to\infty} (f_{n_k}, g) = (f, g).$$

One says that $\{f_{n_k}\}$ **converges weakly** to f.

Solution. The proof is similar to the Arzela-Ascoli theorem, and as with that proof, the main hang-up is notation. Let $\{e_n\}$ be an orthonormal basis for $\mathcal{H}$. (We're assuming here that $\mathcal{H}$ is separable, of course, which the book includes in its definition of Hilbert space.) Then we can define f by defining its inner product with each e_n. Now $\langle f_n, e_1\rangle$ is a sequence of real numbers in $[-1, 1]$ (by the Cauchy-Schwarz inequality, since these are both unit vectors). Since $[-1, 1]$ is compact, there is a subsequence f_{n_j} which converges to some limit ℓ_1. Then $\langle f_{n_j}, e_2\rangle$ is a sequence of real numbers in $[-1, 1]$, so some subsequence $f_{n_{j_r}}$ converges to a limit ℓ_2. Continuing in this fashion, we obtain sequences $S_0, S_1, S_2, \ldots$ where S_{m+1} is a subsequence of S_m, $S_0 = \{f_n\}$, and $\langle v_n^{(m)}, e_j\rangle \xrightarrow{n} \ell_j$ for $j = 1, \ldots, m$, where $v_n^{(m)}$ is the nth term of S_m. Define a sequence S whose kth term is $v_k^{(k)}$. Then for any S_m, the tail of S is a subsequence of the tail of S_m. Hence $\langle v_k, e_j\rangle \xrightarrow{k} \ell_j$ for all j. Define f by $\langle f, e_j\rangle = \ell_j$. Relabeling, $v_k = f_{n_k}$ is a subsequence with $\langle f_{n_k}, e_j\rangle \to \langle f, e_j\rangle$ for all basis elements e_j, and thus $\langle f_{n_k}, g\rangle \to \langle f, g\rangle$ for all $g \in \mathcal{H}$. □

Exercise 17: Fatou's theorem can be generalized by allowing a point to approach the boundary in larger regions, as follows.

For each $0 < s < 1$ and point z on the unit circle, consider the region $\Gamma_s(z)$ defined as the smallest closed convex set that contains z and the closed disc $D_s(0)$. In other words, $\Gamma_s(z)$ consists of all lines joining z with

pointsn in $D_s(0)$. Near the point z, the region $\Gamma_s(z)$ looks like a triangle. See Figure 2.

We say that a function F defined in the open unit disc has a **non-tangential limit** at a point z on the circle, if for every $0 < s < 1$, the limit

$$\underset{\substack{w \to z \\ w \in \Gamma_s(z)}}{F(w)}$$

exists.

Prove that if F is holomorphic and bounded on the open unit disc, then F has a non-tangential limit for almost every point on the unit circle.

Solution. Since F is holomorphic, we have $F(z) = \sum_{n=0}^{\infty} a_n z^n$ for $|z| < 1$. As shown on page 174 in the proof of Fatou's theorem, $\sum |a_n|^2 < \infty$ so there is an $L^2(T)$ function $F(e^{i\theta})$ whose Fourier coefficients are a_n. Note also that $F(e^{i\theta})$ is bounded (almost everywhere) since, by Fatou's theorem, it is the a.e. radial limit of $F(z)$, so $|F(z)| \leq M \Rightarrow |F(e^{i\theta})| \leq M$.

We next prove a lemma about the Poisson kernel.

Lemma 2. *For each $s \in (0,1)$ there exists a constant k_s such that*

$$P_r(\theta - \phi) \leq k_s P_r(-\phi)$$

for all (r, θ) such that $re^{i\theta} \in \Gamma_s$.

Proof. By elementary arithmetic,

$$P_r(\theta - \phi) = \frac{1 - r^2}{|e^{i\phi} - re^{i\theta}|^2}$$

and

$$P_r(-\phi) = \frac{1 - r^2}{|e^{i\phi} - r|^2}.$$

(This alternate formula can be found in any complex analysis book.) Our task is thus reduced to proving

$$|e^{i\phi} - r| \leq k_s |e^{i\phi} - re^{i\theta}|.$$

By the triangle inequality,

$$|e^{i\phi} - r| \leq |e^{i\phi} - re^{i\theta}| + r|e^{i\theta} - 1| = |e^{i\phi} - re^{i\theta}| + 2r \sin\left(\frac{|\theta|}{2}\right).$$

Thus, our task is reduced to proving that

$$\frac{2r \sin\left(\frac{\theta}{2}\right)}{|e^{i\phi} - re^{i\theta}|}$$

is bounded on Γ_s. But $|e^{i\phi} - re^{i\theta}| \geq 1 - r$ by the triangle inequality, so it is sufficient to prove that

$$\frac{2r \sin\left(\frac{\theta}{2}\right)}{1 - r}$$

is bounded on Γ_s. Now for each r, the maximum value of $|\theta|$ (which will maximize this quotient) is, as indicated in the diagram below, one for which $2r \sin\left(\frac{\theta}{2}\right)$ occurs in a triangle with $1 - r$ and $\sqrt{1 - s^2} - \sqrt{r^2 - s^2}$. Thus,

it is at most equal to the sum of them (actually quite a bit less). So we finally need only to prove that

$$\frac{\sqrt{1-s^2}-\sqrt{r^2-s^2}}{1-r}$$

is bounded as $r \to 1$. But this follows from the fact that it has negative derivative, since

$$\frac{d}{dr}\left(\sqrt{1-s^2}-\sqrt{r^2-s^2}\right) = -\frac{r}{\sqrt{r^2-s^2}} < -1 = \frac{d}{dr}(1-r).$$

This completes the proof of the lemma. (I know, there are probably much shorter proofs, but it's late at night...) $\square$

Having established this lemma, the rest of the problem becomes trivial. By the Poisson integral formula,

$$\begin{aligned}|F(re^{i\theta})| &= \left|\frac{1}{2\pi}\int_{-\pi}^{\pi} P_r(\theta-\phi)F(e^{i\phi}d\phi\right| \\ &\leq \frac{1}{2\pi}\int_{-\pi}^{\pi} P_r(\theta-\phi)|F(e^{i\phi}|d\phi \\ &\leq k_s\frac{1}{2\pi}\int_{-\pi}^{\pi} P_r(-\phi)|F(e^{i\phi}|d\phi\end{aligned}$$

and by Math 245A (specifically, the fact that P_r is an approximate identity) this last integral tends to zero. (Recall that we assumed $F(1) = 0$.) □

Exercise 21: There are several senses in which a sequence of bounded operators $\{T_n\}$ can converge to a bounded operator T (in a Hilbert space $\mathcal{H}$). First, there is convergence in the norm, that is, $\|T_n - T\| \to 0$ as $n \to \infty$. Next, there is a weaker convergence, which happens to be called **strong convergence**, that requires that $T_n f \to Tf$ as $n \to \infty$ for every $f \in \mathcal{H}$. Finally, there is **weak convergence** (see also Exercise 20) that requires $(T_n f, g) \to (Tf, g)$ for every pair of vectors $f, g \in \mathcal{H}$.

(a) Show by examples that weak convergence does not imply strong convergence, nor does strong convergence imply convergence in the norm.

(b) Show that for any bounded operator T there is a sequence $\{T_n\}$ of bounded operators of finite rank so that $T_n \to T$ strongly as $n \to \infty$.

Solution. (a) Let $\mathcal{H} = \ell^2(\mathbb{N})$. Let T be the zero operator and $T_n = R^n$ where R is the right shift operator; thus

$$T_n(a_1, a_2, \dots) = (0, \dots, 0, a_1, a_2, \dots).$$

Then for any fixed $f = (a_1, a_2, \dots)$ and $g = (b_1, b_2, \dots)$,

$$\langle T_n f, g\rangle = \sum_{k=1}^{\infty} a_k b_{n+k} = \langle f, L^n g\rangle$$

where L is the left-shift operator defined by $Lg = (b_2, b_3, \dots)$. By the Cauchy-Schwarz inequality,

$$|\langle T_n f, g\rangle| \leq \|f\|\|L^n g\|.$$

But $\|L^n g\| \to 0$ because $\sum |b_k|^2$ converges and the tails of a convergent series tend to zero. Hence $\langle T_n f, g\rangle \to 0$ for all $f, g \in \mathcal{H}$. Thus, $T_n \to T$ weakly. However, T_n does not converge to T strongly; note that $\|Rf\| = \|f\|$ for any f, so $\|T_n f\| = \|R^n f\| = \|f\|$ by induction. Hence $T_n f \not\to Tf = 0$.

To show that strong convergence does not imply convergence in norm, let $T_n = L^n$ in the same space. Then for any f, $\|T_n f\| = \|L^n f\| \to 0$ because the tails of a convergent series tend to zero. Hence $T_n \to T$ strongly. However, $\|T_n\| = 1$ because a unit vector $(0, \dots, 0, 1, 0, \dots)$ with more than n initial zeros is mapped to another unit vector. Hence $\|T_n - T\| = \|T_n\| \not\to 0$.

(b) (We assume $\mathcal{H}$ is separable.) Let $\{e_i\}$ be an orthonormal basis for $\mathcal{H}$. We can write

$$Te_i = \sum_{j=1}^{\infty} c_{ij} e_j$$

for each $i = 1, 2, \ldots$. Define

$$T_n e_i = \sum_{j=1}^{n} c_{ij} e_j$$

and extend linearly from the basis to the rest of the space (actually, extend linearly to finite linear combinations of the basis, and then take limits to get the rest of the space...) Clearly each T_n is of finite rank since its range is spanned by $e_1, \ldots, e_n$. Now let $f = \sum_{i=1}^{\infty} a_i e_i$. Then

$$T_f = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_i c_{ij} e_j$$

whereas

$$T_n f = \sum_{i=1}^{\infty} \sum_{j=1}^{n} a_i c_{ij} e_j$$

which is just the nth partial sum (in j) and hence converges to Tf. (This is where we use the fact that T is a bounded operator, since absolute convergence allows us to rearrange these sums.) Hence $T_n f \to Tf$ weakly for all $f \in \mathcal{H}$.

□

Exercise 22: An operator T is an **isometry** if $\|Tf\| = \|f\|$ for all $f \in \mathcal{H}$.

(a) Show that if T is an isometry, then $(Tf, Tg) = (f, g)$ for every $f, g \in \mathcal{H}$. Prove as a result that $T^*T = I$.
(b) If T is an isometry and T is surjective, then T is unitary and $TT^* = I$.
(c) Give an example of an isometry that is not unitary.
(d) Show that if T^*T is unitary then T is an isometry.

Solution. (a) By the polarization identity,

$$\begin{aligned}
\langle Tf, Tg \rangle &= \frac{\|Tf + Tg\|^2 - \|Tf - Tg\|^2 + i\|Tf + iTg\|^2 - i\|Tf - iTg\|^2}{4} \\
&= \frac{\|T(f+g)\|^2 - \|T(f-g)\|^2 + i\|T(f+ig)\|^2 - i\|T(f-ig)\|^2}{4} \\
&= \frac{\|f+g\|^2 - \|f-g\|^2 + i\|f+ig\|^2 - i\|f-ig\|^2}{4} \\
&= \langle f, g \rangle.
\end{aligned}$$

This in turn implies

$$\langle f, T^*Tg \rangle = \langle f, Ig \rangle$$

for all f, g, so that $T^*T = I$.

(b) T preserves norms because it's an isometry; it's injective because norm-preserving linear maps are always injective (since the kernel cannot contain anything nonzero). Since it's surjective as well, it's a norm-preserving linear bijection, which is by definition a unitary map. We

know $T^*T = I$ from part (a). Since T is bijective, it has a linear 2-sided inverse, and the equation $T^*T = 1$ shows that T is this inverse. Hence $TT^* = I$.

(c) The right-shift operator on $\ell^2(\mathbb{N})$ is isometric, since

$$\|(0, a_1, a_2, \dots)\|^2 = \sum_{j=1}^{\infty} |a_j|^2 = \|(a_1, a_2, \dots)\|^2.$$

However, this operator is not surjective, so it's not unitary.

(d) First,

$$\|Tf\|^2 = \langle Tf, Tf\rangle = \langle f, T^*Tf\rangle \leq \|f\|\|T^*Tf\| = \|f\|^2$$

by the Cauchy-Schwarz inequality; hence $\|Tf\| \leq \|f\|$. Then,

$$\|f\|^2 = \|T^*Tf\|^2 = \langle T^*Tf, T^*Tf\rangle = \langle Tf, TT^*Tf\rangle \leq \|Tf\|\|T(T^*Tf)\| \leq \|Tf\|\|T^*Tf\| = \|Tf\|\|f\|$$

where we have applied the previous inequality with T^*Tf in place of f. Dividing by $\|f\|$ yields $\|f\| \leq \|Tf\|$. Putting the inequalities together, $\|Tf\| = \|f\|$. Hence T is an isometry.

□

Exercise 23: Suppose $\{T_k\}$ is a collection of bounded operators on a Hilbert space $\mathcal{H}$, with $\|T_k\| \leq 1$ for all k. Suppose also that

$$T_k T_j^* = T_k^* T_j = 0 \quad \text{for all } k \neq j.$$

Let $S_N = \sum_{k=-N}^{n} T_k$. Show that $S_N(f)$ converges as $N \to \infty$, for every $f \in \mathcal{H}$. If $T(f)$ denotes the limite, prove that $\|T\| \leq 1$.

Exercise 24: Let $\{e_k\}_{k=1}^{\infty}$ denote an orthonormal set in a Hilbert space $\mathcal{H}$. If $\{c_k\}_{k=1}^{\infty}$ is a sequence of positive real numbers such that $\sum c_k^2 < \infty$, then the set

$$A = \left\{ \sum_{k=1}^{\infty} a_k e_k : |a_k| \leq c_k \right\}$$

is compact in $\mathcal{H}$.

Solution. This is a standard diagonalization argument. Let Σ be a sequence of points in $\mathcal{H}$. Consider the first components of the points in Σ, i.e. their components with respect to e_1. This is a sequence of complex numbers in the compact ball $\{z : |z| \leq c_1\}$, so some subsequence converges to a complex number in this ball. If we take the corresponding subsequence of Σ, we obtain a subsequence S_1 whose first components converge to some number b_1 with $|b_1| \leq c_1$. Now consider the second components of the points in S_1; they form a sequence of complex numbers in the compact ball $\{z : |z| \leq c_2\}$ and hence some subsequence converges to a complex number b_2 in this ball. Taking the corresponding subsequence S_2 of S_1, we have a sequence whose first and second components converge. Continuing, we may inductively define sequences S_n for all n such that S_{n+1} is a subsequence of S_n, and the first n components of S_n converge. Finally, we define a sequence S whose nth term is the nth term of S_n. This is a subsequence of Σ, and for any n it is eventually a subsequence of S_n (i.e. the tail of S is a subsequence of the tail of S_n). Hence it converges in every component.

But this implies convergence to a point in the Hilbert space (since the sizes of the tails are uniformly bounded), so we are done. □

Exercise 25: Suppose T is a bounded operator that is diagonal with respect to a basis $\{\phi_k\}$, with $T\phi_k = \lambda_k \phi_k$. Then T is compact if and only if $\lambda_k \to 0$.

Solution. (From lecture) Suppose $\lambda_k \to 0$. Let T_n be the nth truncation, i.e. the operator that results when λ_k is replaced with 0 for $k > n$. Then $T - T_n$ is also a diagonal operator, with

$$\|T - T_n\| = \sup_{k>n} |\lambda_k| \to 0.$$

Since T can be uniformly approximated by operators of finite rank, it is compact. Conversely, suppose that $\lambda_k \not\to 0$, i.e. $\limsup |\lambda_n| > 0$, so that there is some subsequence λ_{n_j} with $|\lambda_{n_j}| > \frac{\delta}{2}$ for some real number $\delta > 0$. Then $T\phi_{n_j} = \lambda_{n_j}\phi_{n_j}$ and by orthonormality,

$$\|T\phi_{n_j} - T\phi_{n_k}\| = \sqrt{\lambda_{n_j}^2 + \lambda_{n_k}^2} > \frac{\delta}{\sqrt{2}}.$$

Since all the points of the sequence $\{T\phi_{n_j}\}$ are uniformly bounded away from each other, it can have no convergent subsequence. These points all lie in $\overline{T(B)}$, so $T(B)$ is not compact. □

Exercise 26: Suppose w is a measurable function on $\mathbb{R}^d$ with $0 < w(x) < \infty$ for a.e. x, and K is a measurable function on $\mathbb{R}^{2d}$ that satisfies:

(i)

$$\int_{\mathbb{R}^d} |K(x,y)| w(y) dy \le A w(x) \quad \text{for almost every } x \in \mathbb{R}^d, \text{ and}$$

(ii)

$$\int_{\mathbb{R}^d} |K(x,y)| w(x) dx \le A w(y) \quad \text{for almost every } y \in \mathbb{R}^d.$$

Prove that the integral operator defined by

$$Tf(x) = \int_{\mathbb{R}^d} K(x,y) f(y) dy, \quad x \in \mathbb{R}^d$$

is bounded on $L^2(\mathbb{R}^d)$ with $\|T\| \le A$. Note as a special case that if $\int |K(x,y)| dy \le A$ for all x, and $\int |K(x,y)| dx \le A$ for all y, then $\|T\| \le A$.

Solution. First, note that

$$\begin{aligned}
\left(\int |K(x,y)||f(y)| dy \right)^2 &= \left(\int \left(\sqrt{|K(x,y)|}\sqrt{w(y)} \right) \left(\sqrt{|K(x,y)|}|f(y)|\sqrt{w(y)}^{-1} \right) \right)^2 \\
&\le \left(\int |K(x,y)| w(y) dy \right) \left(\int |K(x,y)||f(y)|^2 w(y)^{-1} dy \right) \\
&\overset{(a.e.)}{\le} A w(x) \int |K(x,y)||f(y)|^2 w(y)^{-1} dy
\end{aligned}$$

by the Cauchy-Schwarz inequality. Thus, for $f \in L^2$,

$$\begin{aligned}
\|Tf\|^2 &= \int \left| \int K(x,y)f(y)dy \right|^2 dx \\
&\le \int \left(\int |K(x,y)||f(y)|dy \right)^2 dx \\
&\le \int Aw(x) \int |K(x,y)||f(y)|^2 w(y)^{-1} dydx \\
&\overset{(Tonelli)}{=} A \int |f(y)|^2 w(y)^{-1} \int |K(x,y)|w(x)dxdy \\
&\le A \int |f(y)|^2 w(y)^{-1} Aw(y)dy \\
&= A^2 \int |f(y)|^2 dy \\
&= A^2 \|f\|^2.
\end{aligned}$$

Hence $\|T\| \le A$. □

Exercise 27: Prove that the operator

$$Tf(x) = \frac{1}{\pi} \int_0^\infty \frac{f(y)}{x+y} dy$$

is bounded on $L^2(0,\infty)$ with norm $\|T\| \le 1$.

Exercise 28: Suppose $\mathcal{H} = L^2(B)$, where B is the unit ball in $\mathbb{R}^d$. Let $K(x,y)$ be a measurable function on $B \times B$ that satisfies $|K(x,y)| \le A|x-y|^{-d+\alpha}$ for some $\alpha > 0$, whenever $x, y \in B$. Define

$$Tf(x) = \int_B K(x,y)f(y)dy.$$

(a) Prove that T is a bounded operator on $\mathcal{H}$.
(b) Prove that T is compact.
(c) Note that T is a Hilbert-Schmidt operator if and only if $\alpha > d/2$.

Solution.

(a) Let

$$C = \int\limits_{z \in \mathbb{R}^d : |z| \le 2} \frac{dz}{|z|^{d-\alpha}}$$

which converges because the exponent is less than d. Then

$$\int_B |K(x,y)|dy \le \int \frac{Ady}{|x-y|^{d-\alpha}} \le A \int_{|z|\le 2} \frac{dz}{|z|^{d-\alpha}} = AC$$

so by problem 26 with $w = 1$, we have T bounded with $\|T\| \le AC$.

(b) As suggested, let

$$K_n(x,y) = \begin{cases} K(x,y) & |x-y| \ge \frac{1}{n} \\ 0 & \text{else} \end{cases}$$

and

$$T_n f(x) = \int K_n(x,y)f(y)dy.$$

Then T_n is Hilbert-Schmidt (and therefore compact) since clearly $K_n \in L^2(B\times B)$ (K_n is, after all, bounded with compact support). Moreover,

$$\int_B |K_n(x,y) - K(x,y)|dy \leq \int_{|x-y|\leq 1/n} A|x-y|^{-d+\alpha} = AC_n$$

where we define

$$C_n = \int\limits_{z\in\mathbb{R}^d:|z|\leq 1/n} \frac{dz}{|z|^{d-\alpha}}.$$

Since $\frac{1}{|z|^{d-\alpha}} \in L^1(\mathbb{R}^d)$, the absolute continuity of the integral implies $C_n \to 0$. By problem 26 again with $w = 1$, this implies $\|T - T_n\| \to 0$. Since T_n is compact, this implies that T is compact.

(c) This should actually say "T is *guaranteed* to be Hilbert-Schmidt if and only if..." since K could be **a lot** less than the bound given. Anyhoo,

$$\begin{aligned} T \text{ necessarily Hilbert-Schmidt} &\Leftrightarrow A|x-y|^{-d+\alpha} \in L^2(B\times B) \\ &\Leftrightarrow \int_B\int_B A^2|x-y|^{-2d+2\alpha} < \infty \\ &\Leftrightarrow -2d+2\alpha > -d \\ &\Leftrightarrow \alpha > \frac{d}{2}. \end{aligned}$$

□

Exercise 29: Let T be a compact operator on a Hilbert space $\mathcal{H}$ and assume $\lambda \neq 0$.

(a) Show that the range of $\lambda I - T$ defined by

$$\{g \in \mathcal{H} : g = (\lambda I - T)f \text{ for some } f \in \mathcal{H}\}$$

is closed.

(b) Show by example that this may fail when $\lambda = 0$.

(c) Show that the range of $\lambda I - T$ is all of $\mathcal{H}$ if and only if the null space of $\bar{\lambda} I - T^*$ is trivial.

Exercise 30: Let $\mathcal{H} = L^2([-\pi,\pi])$ with $[-\pi,\pi]$ identified as the unit circle. Fix a bounded sequence $\{\lambda_n\}_{n=-\infty}^{\infty}$ of complex numbers, and define an operator Tf by

$$Tf(x) \sim \sum_{n=-\infty}^{\infty} \lambda_n a_n e^{inx} \quad \text{whenever} \quad f(x) \sim \sum_{n=-\infty}^{\infty} a_n e^{inx}.$$

Such an operator is called a **Fourier multiplication operator**, and the sequence $\{\lambda_n\}$ is called the **multiplier sequence**.

(a) Show that T is a bounded operator on $\mathcal{H}$ and $\|T\| = \sup|\lambda_n|$.

(b) Verify that T commutes with translations, that is, if we define $\tau_h(x) = f(x-h)$ then

$$T \circ \tau_h = \tau_h \circ T \quad \text{for every } h \in \mathbb{R}.$$

(c) Conversely, prove that if T is any bounded operator on $\mathcal{H}$ that commutes with translations, then T is a Fourier multiplier operator.

Exercise 34: Let K be a Hilbert-Schmidt kernel which is real and symmetric. Then, as we saw, the operator T whose kernel is K is compact and symmetric. Let $\{\phi_k(x)\}$ be the eigenvectors (with eigenvalues λ_k) that diagonalize T. Then

(a) $\sum |\lambda_k|^2 < \infty$.

(b) $K(x,y) \sim \sum \lambda_k \phi_k(x)\phi_k(y)$ is the expansion of K in the basis $\{phi_k(x)\phi_k(y)\}$.

(c) Suppose T is a compact operator which is symmetric. Then T is of Hilbert-Schmidt type if and only if $\sum |\lambda_n|^2 < \infty$, where $\{\lambda_n\}$ are the eigenvalues of T counted according to their multiplicities.

Exercise 35: Let $\mathcal{H}$ be a Hilbert space. Prove the following variants of the spectral theorem.

(a) If T_1 and T_2 are two linear symmetric and compact operators on $\mathcal{H}$ that commute, show that they can be diagonalized simultaneously. In other words, there exists an orthonormal basis for $\mathcal{H}$ which consists of eigenvectors for both T_1 and T_2.

(b) A linear operator on $\mathcal{H}$ is **normal** if $TT^* = T^*T$. Prove that if T is normal and compact, then T can be diagonalized.

(c) If U is unitary, and $U = \lambda I - T$ where T is compact, then U can be diagonalized.

Solution.

(a) We can pretty much copy the proof verbatim with "eigenvector" replaced by "common eigenvector". Let S be the closure of the subspace of $\mathcal{H}$ spanned by all common eigenvectors of T_1 and T_2. We want to show $S = \mathcal{H}$. Suppose not; then $\mathcal{H} = S \oplus S^\perp$ with $S^\perp$ nonempty. If we can show $S^\perp$ contains a common eigenvector of T_1 and T_2, we have a contradiction. Note that $T_1 S \subset S$, which in turn implies $T_1 S^\perp \subset S^\perp$ since

$$g \in S^\perp \Rightarrow \langle Tg, f\rangle = \langle g, Tf\rangle = 0$$

for all $f \in S$. Similarly, $T_2 S^\perp \subset S^\perp$. Now by the theorem for one operator, T_1 must have an eigenvector in $S^\perp$ with some eigenvalue λ. Let E_λ be the eigenspace of λ (as a subspace of $S^\perp$). Then for any $x \in E_\lambda$,

$$T_1(T_2 x) = T_2(T_1 x) = T_2(\lambda x) = \lambda(T_2 x)$$

so $T_2 x \in E_\lambda$ as well. Since T_2 fixes E_λ, it has at least one eigenvector in E_λ. This eigenvector is a common eigenvector of T_1 and T_2, providing us with our contradiction.

(b) This follows from part (a). Write

$$T = \frac{T+T^*}{2} + i\frac{T-T^*}{2i}.$$

By a trivial calculation, both $\frac{T+T^*}{2}$ and $\frac{T-T^*}{2i}$ are self-adjoint. Moreover, since T is normal,

$$(T+T^*)(T-T^*) = T^2 + T^*T - TT^* - T^{*2} = T^2 - T^{*2} = (T-T^*)(T+T^*)$$

so they commute as well. Hence, there exists an ONB of common eigenvectors of $\frac{T+T^*}{2}$ and $\frac{T-T^*}{2i}$. Any such common eigenvector is an

eigenvector of T, since

$$\frac{T+T^*}{2}x = \lambda x \text{ and } \frac{T-T^*}{2i}x = \lambda' x \Rightarrow Tx = (\lambda + i\lambda')x.$$

(c)

□

Chapter 4.8, Page 202

Problem 1: Let $\mathcal{H}$ be an infinite-dimensional Hilbert space. There exists a linear functional ℓ defined on $\mathcal{H}$ that is not bounded.

Solution. It is a well-known fact from linear algebra that every vector space has a basis. This can be proved using Zorn's lemma: linearly independent sets are partially ordered by inclusion, and every chain has an upper bound by union, so there exists a maximal linearly independent set, which is by definition a basis. Applying this to our Hilbert space, we obtain an (algebraic) basis, i.e. one for which every vector is a *finite* linear combination of basis elements. Let $\{e_n\}$ be a countable subset of our algebraic basis. Define $\ell(e_n) = n\|e_n\|$ and $\ell(f) = 0$ for f in our basis but $f \neq e_n$ for any n. We can then extend ℓ to the whole space in a well-defined manner, but clearly ℓ is not bounded since $|\ell(e_n)| = n\|e_n\|$. □

Problem 2: The following is an example of a non-separable Hilbert space. We consider the collection of exponentials $\{e^{i\lambda x}\}$ on $\mathbb{R}$, where λ ranges over the real numbers. Let $\mathcal{H}_0$ denote the space of finite linear combinations of these exponentials. For $f, g \in \mathcal{H}_0$, we define the inner product as

$$(f,g) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} f(x)\overline{g(x)}dx.$$

(a) Show that this limit exists, and

$$(f,g) = \sum_{k=1}^{N} a_{\lambda_k} \bar{b_{\lambda_k}}$$

if $f(x) = \sum_{k=1}^N a_{\lambda_k} e^{i\lambda_k x}$ and $g(x) = \sum_{k=1}^N b_{\lambda_k} e^{i\lambda_k x}$.

(b) With this inner product $\mathcal{H}_0$ is a pre-Hilbert space. Notice that $\|f\| \le \sup_x |f(x)|$, if $f \in \mathcal{H}_0$, where $\|f\|$ denotes the norm $\langle f, f\rangle^{1/2}$. Let $\mathcal{H}$ be the completion of $\mathcal{H}_0$. Then $\mathcal{H}$ is not separable because $e^{i\lambda x}$ and $e^{i\lambda' x}$ are orthonormal if $\lambda \neq \lambda'$. A continuous function F defined on $\mathbb{R}$ is called **almost periodic** if it is the uniform limit (on $\mathbb{R}$) of elements in $\mathcal{H}_0$. Such functions can be identified with (certain) elements in the completion $\mathcal{H}$: We have $\mathcal{H}_0 \subset AP \subset \mathcal{H}$, where AP denotes the almost periodic functions.

(c) A continuous function F is in AP if for ever $\epsilon > 0$ we can find a length $L = L_\epsilon$ such that any interval $I \subset \mathbb{R}$ of length L contains an "almost period" τ satisfying

$$\sup_x |F(x+\tau) - F(x)| < \epsilon.$$

(d) An equivalent characterization is that F is in AP if and only if every sequence $F(x+h_n)$ of translates of F contains a subsequence that converges uniformly.

Problem 7: Show that the identity operator on $L^2(\mathbb{R}^d)$ cannot be given as an (absolutely) convergent integral operator. More precisely, if $K(x,y)$ is a measurable function on $\mathbb{R}^d \times \mathbb{R}^d$ with the property that for each $f \in L^2(\mathbb{R}^d)$, the integral $T(f)(x) = \int_{\mathbb{R}^d} K(x,y)f(y)dy$ converges for almost every x, then $T(f) \neq f$ for some f.

Solution. Suppose such a K exists. Let B_1 and B_2 be disjoint balls in $\mathbb{R}^d$. We will show that $K = 0$ a.e. in $B_1 \times B_2$. Suppose not; then there is a "rectangle" $E_1 \times E_2$ with $E_1 \subset B_1$ and $E_2 \subset B_2$ sets of positive measure, on which $K(x,y) > 0$ or $K(x,y) < 0$; WLOG, $K(x,y) > 0$. (If K is allowed complex values, we can change this condition to $\mathrm{Re}(K) > 0$.) Let $f = \chi_{E_2}$. Then for almost all $x \in E_1$,

$$0 = f(x) = Tf(x) = \int_{E_2} K(x,y)dy.$$

But the integral of a positive function over a set of positive measure is nonzero, so we have a contradiction. Thus, $K(x,y) = 0$ a.e. in $B_1 \times B_2$. Now if we let $\Delta = \{(x,x) : x \in \mathbb{R}^d\} \subset \mathbb{R}^{2d}$ be the "diagonal", we can cover $\mathbb{R}^{2d} \setminus \Delta$ with product sets $B_1 \times B_2$. (One way to see this is topological: products of balls form a basis for the product topology on $\mathbb{R}^d \times \mathbb{R}^d$, and Δ is closed, so its complement is open.) This implies $K(x,y) = 0$ a.e. on $\mathbb{R}^{2d} \setminus \Delta$, and therefore a.e. on $\mathbb{R}^{2d}$. But then $Tf = 0$ for all f, so any nonzero f will have $Tf \neq f$. □

Problem 8: Suppose $\{T_k\}$ is a collection of bounded operators on a Hilbert space $\mathcal{H}$. Assume that

$$\|T_k T_j^*\| \leq a_{k-j}^2 \quad \text{and} \quad \|T_k^* T_j\| \leq a_{k-j}^2,$$

for positive constants $\{a_n\}$ with the property that $\sum_{n=-\infty}^{\infty} a_n = A < \infty$. Then $S_N(f)$ converges as $N \to \infty$, for every $f \in \mathcal{H}$, with $S_N = \sum_{-N}^{N} T_k$. Moreover, $T = \lim_{N\to\infty} S_N$ satisfies $\|T\| \leq A$.

Solution. For any integers N and n,

$$\begin{aligned} \|S_N\|^{2^n} &= \|(S^*S)^{2^{n-1}}\| \\ &= \left\| \sum_{j_1=-N}^{N} \sum_{k_1=-N}^{N} \cdots \sum_{j_{2^{n-1}}=-N}^{N} \sum_{k_{2^{n-1}}=-N}^{N} T_{j_1}^* T_{k_1} \cdots T_{j_{2^{n-1}}}^* T_{k_{2^{n-1}}} \right\| \\ &\leq \sum_{j_1=-N}^{N} \sum_{k_1=-N}^{N} \cdots \sum_{j_{2^{n-1}}=-N}^{N} \sum_{k_{2^{n-1}}=-N}^{N} \|T_{j_1}^* T_{k_1} \cdots T_{j_{2^{n-1}}}^* T_{k_{2^{n-1}}}\|. \end{aligned}$$

Now since $\|T_m\| \leq a_0$ for any m,

□

Problem 9: A discussion of a class of regular Sturm-Liouville operators follows. Other special examples are given in the problems below.

Suppose $[a,b]$ is a bounded interval, and L is defined on functions f that are twice continuously differentiable in $[a,b]$ (we write $f \in C^2([a,b])$ by

$$L(f)(x) = \frac{d^2 f}{dx^2} - q(x)f(x).$$

Here the function q is continuous and real-valued on $[a,b]$, and we assume for simplicity that q is non-negative. We say that $\phi \in C^2([a,b])$ is an **eigenfunction** of L with eigenvalue μ if $L(\phi) = \mu\phi$, under the assumption that ϕ satisfies the boundary conditions $\phi(a) = \phi(b) = 0$. Then one can show:

(a) The eigenvalues μ are strictly negative, and the eigenspace corresponding to each eigenvalue is one-dimensional.

(b) Eigenvectors corresponding to distinct eigenvalues are orthogonal in $L^2([a,b])$.

(c) Let $K(x,y)$ be the "Green's kernel" defined as follows. Choose $\phi_-(x)$ to be a solution of $L(\phi_-) = 0$, with $\phi_-(a) = 0$ but $\phi_-'(a) \neq 0$. Similarly, choose $\phi_+(x)$ to be a solution of $L(\phi_+) = 0$ with $\phi_+(b) = 0$ but $\phi_+'(b) \neq 0$. Let $w = \phi_+'(x)\phi_-(x) - \phi_-'(x)\phi_+(x)$, be the "Wronskian" of these solutions, and note that w is a non-zero constant.
Set

$$K(x,y) = \begin{cases} \frac{\phi_-(x)\phi_+(y)}{w} & a \le x \le y \le b, \\ \frac{\phi_+(x)\phi_-(y)}{w} & a \le y \le x \le b. \end{cases}$$

Then the operator T defined by

$$T(f)(x) = \int_a^b K(x,y)f(y)dy$$

is a Hilbert-Schmidt operator, and hence compact. It is also symmetric. Moreover, whenever f is continuous on $[a,b]$, Tf is of class $C^2([a,b])$ and

$$L(Tf) = f.$$

(d) As a result, each eigenvector of T (with eigenvalue λ) is an eigenvector of L (with eigenvalue $\mu = 1/\lambda$). Hence Theorem 6.2 proves the completeness of the orthonormal set arising from normalizing the eigenvectors of L.

Solution.

(a) Let ϕ be an eigenfunction of L with eigenvalue μ. Then

$$\phi'' = (q+\mu)\phi \Rightarrow \phi\phi'' = (q+\mu)\phi^2.$$

Integrating by parts from a to b, we have

$$\phi\phi'|_a^b - \int (\phi')^2 = \int (q+\mu)\phi^2.$$

Since $\phi(a) = \phi(b) = 0$, this reduces to

$$-\int (\phi')^2 = \int (q+\mu)\phi^2.$$

Now if ϕ is not almost everywhere zero, the LHS is strictly negative. But the integrand on the RHS is everywhere nonnegative unless μ is strictly negative. Hence the eigenvalues of L are all strictly negative. Now suppose μ is an eigenvalue of L with eigenfunctions ϕ_1 and ϕ_2. Then

$$\begin{aligned}(\mu+q)\phi_1\phi_2 = \phi_1\phi_2'' &= \phi_2\phi_1'' \\ &\Rightarrow \phi_1'\phi_2' + \phi_1\phi_2'' = \phi_1'\phi_2' + \phi_1''\phi_2 \\ &\Rightarrow (\phi_1\phi_2')' = (\phi_1'\phi_2)' \\ &\Rightarrow \phi_1\phi_2' = \phi_1'\phi_2 + C.\end{aligned}$$

Plugging in a, we see that $C = 0$ because $\phi_1(a) = \phi_2(a) = 0$. Hence $\phi_1\phi_2' - \phi_1'\phi_2 = 0$. Since the Wronskian of these solutions is zero, they are linearly dependent. So the eigenspace can only be one-dimensional.

(b) Suppose ϕ_1 and ϕ_2 are eigenfunctions with eigenvalues μ_1 and μ_2 respectively. Then

$$\begin{aligned}\mu_1 \int \phi\psi &= \int (\phi'' - q\phi)\psi \\ &= \int \phi''\psi - \int q\phi\psi \\ &= \phi'\psi|_a^b - \int \phi'\psi' - \int q\phi\psi \\ &= -\int \phi'\psi' - \int q\phi\psi \\ &= \phi\psi'|_a^b - \int \phi'\psi' - \int q\phi\psi \\ &= \int \psi''\phi - \int q\phi\psi \\ &= \int (\psi'' - q\psi)\phi \\ &= \mu_2 \int \psi\phi.\end{aligned}$$

If $\mu_1 \neq \mu_2$, this implies $\int \phi_1\phi_2 = 0$.

(c) The Wronskian is a constant because

$$w' = \phi_+''\phi_- - \phi_-''\phi_+ = q\phi_+\phi_- - q\phi_-\phi_+ = 0;$$

it is nonzero because plugging in at a yields $-\phi_-'(a)\phi_+(a)$. We already know $\phi_-'(a) \neq 0$, and $\phi_+(a)$ cannot be zero because then ϕ_+ would be an eigenfunction with eigenvalue 0, contradicting part (a). To show T is Hilbert-Schmidt, consider $\int\int |K^2|$. We will treat this on the region $R = \{a \leq y \leq x \leq b\}$; the other half is symmetric. Then

$$\iint_R |K(x,y)|^2 dxdy = \iint_R \frac{|\phi_+(x)\phi_-(y)|^2}{w^2} dxdy.$$

Now w is a nonzero constant, as we saw above; ϕ_+ and ϕ_- are both continuous on a compact set and hence bounded. Thus, the integrand is bounded, and the region of integration is compact, so the integral

is finite. So T is Hilbert-Schmidt. The symmetry of T is immediately evident from its definition. Now suppose $f \in C([a,b])$. Then $Tf \in C^2([a,b])$ because $K \in C^2([a,b]^2)$ and the second partials of K are bounded so that one can differentiate Tf under the integral sign. Finally,

$$\begin{aligned} Tf(x) &= \int_a^b K(x,y)f(y)dy = \frac{1}{w}\int_a^x \phi_+(x)\phi_-(y)f(y)dy + \frac{1}{w}\int_x^b \phi_-(x)\phi_+(y)f(y)dy \\ &= \frac{\phi_+(x)}{w}\int_a^x \phi_-(y)f(y)dy + \frac{\phi_-(x)}{w}\int_x^b \phi_+(y)f(y)dy, \end{aligned}$$

so

$$\begin{aligned} (Tf)'(x) &= \frac{\phi_+'(x)}{w}\int_a^x \phi_-(y)f(y)dy + \frac{\phi_+(x)}{w}\phi_-(x)f(x) + \frac{\phi_-'(x)}{w}\int_x^b \phi_+(y)f(y)dy - \frac{\phi_-(x)}{w}\phi_+(x)f(x) \\ &= \frac{\phi_+'(x)}{w}\int_a^x \phi_-(y)f(y)dy + \frac{\phi_-'(x)}{w}\int_x^b \phi_+(y)f(y)dy \end{aligned}$$

and

$$\begin{aligned} (Tf)''(x) &= \frac{\phi_+''(x)}{w}\int_a^x \phi_-(y)f(y)dy + \frac{\phi_+'(x)}{w}\phi_-(x)f(x) + \frac{\phi_-''(x)}{w}\int_x^b \phi_+(y)f(y)dy - \frac{\phi_-'(x)}{w}\phi_+(x)f(x) \\ &= f(x)\frac{w}{w} + \frac{q(x)\phi_+(x)}{w}\int_a^x \phi_-(y)f(y)dy + \frac{q(x)\phi_-(x)}{w}\int_x^b \phi_+(y)f(y)dy \\ &= f(x) + q(x)\int_a^b K(x,y)f(y)dy \\ &= f(x) + q(x)(Tf)(x) \end{aligned}$$

so $L(Tf) = (Tf)'' - q(Tf) = f$.

(d) This is more just an observation in the problem statement than something for me to do.

□

Chapter 5.5, Page 253

Exercise 1: Suppose $f \in L^2(\mathbb{R}^d)$ and $k \in L^1(\mathbb{R}^d)$.

(a) Show that $(f * k)(x) = \int f(x-y)k(y)dy$ converges for a.e. x.
(b) Prove that $\|f * k\|_2 \le \|f\|_2\|k\|_1$.
(c) Establish $\widehat{(f * k)}(\xi) = \hat{k}(\xi)\hat{f}(\xi)$ for a.e. ξ.
(d) The operator $Tf = f * k$ is a Fourier multiplication operator with multiplier $m(\xi) = \hat{k}(\xi)$.

Solution.

(a) This will follow from part (b) because an L^2 function must be finite almost everywhere (which will prove a.e. convergence of $\int |f(x-y)||k(y)|dy$), and absolutely convergent integrals are convergent.
(b) Just so I have it for future reference, why don't I prove the L^p version of this. Suppose $f \in L^p$ with $1 < p < \infty$ and $k \in L^1$. Let q be the

conjugate exponent of p. Then

$$\begin{aligned}
\|f * k\|_p^p &= \int \left| \int f(x-y)k(y)dy \right|^p dx \\
&\leq \int \left(\int |f(x-y)||k(y)|^{1/p}|k(y)|^{1/q}dy \right)^p dx \\
&\leq \int \left(\|f(x-y)k(y)^{1/p}\|_p \|k(y)^{1/q}\|_q \right)^p dx \\
&= \int \left(\int |f(x-y)|^p |k(y)|dy \right) \left(\int |k(y)|dy \right)^{p/q} dx \\
&= \|k\|_1^{p/q} \|f^p * k\|_1 \\
&\leq \|k\|_1^{p/q} \|k\|_1 \|f^p\|_1 \\
&= \|k\|_1^p \|f\|_p^p.
\end{aligned}$$

Here we have used Hölder's inequality on $|f||k|^{1/p} \in L^p$ and $|k|^{1/q} \in L^q$, as well as the bound for the L^1 norm of a convolution of L^1 functions. In the case $p = q = 2$, Hölder's inequality reduces to Cauchy-Schwarz.

(c) If $f \in L^1 \cap L^2$, we already know this from our theory of Fourier transforms on L^1. Otherwise, since $L^1 \cap L^2$ is dense in L^2, we may take a sequence $f_n \in L^1 \cap L^2$ with $f_n \xrightarrow{L^2} f$. Then

$$\begin{aligned}
\left| \widehat{f * k}(\xi) - \widehat{f_n * k}(\xi) \right| &= \left| \int e^{-2\pi i \xi x} \int \left(f(x-y) - f_n(x-y) \right) k(y)dydx \right| \\
&\leq \int \int |f(x-y) - f_n(x-y)| \, |k(y)|dydx \\
&\leq \|f - f_n\|_2 \|k\|_1 \xrightarrow{L^2} 0
\end{aligned}$$

so

$$\widehat{f * k}(\xi) = \lim \widehat{f_n * k}(\xi) = \lim \hat{f}_n(\xi)\hat{k}(\xi) = \hat{k}(\xi) \lim \hat{f}_n(\xi) = \hat{k}(\xi)\hat{f}(\xi).$$

(d) This is just the definition of a Fourier multiplication operator applied to part (c).

□

Exercise 3: Let $F(z)$ be a bounded holomorphic function in the half-plane. Show in two ways that $\lim_{y \to 0} F(x+iy)$ exists for a.e. x.

(a) By using the fact that $F(z)/(z+i)$ is in $H^2(\mathbb{R}^2_+)$.

(b) By noting that $G(z) = F\left(i\frac{1-z}{1+z}\right)$ is a bounded holomorphic function in the unit disc, and using Exercise 17 in the previous chapter.

Solution.

(a) Since $|F(z)| \leq M$ for some M,

$$\left| \frac{F(x+iy)}{x+i(y+1)} \right| \leq \frac{M}{\sqrt{x^2+1}}$$

so

$$\int_{-\infty}^{\infty} \left| \frac{F(x+iy)}{x+i(y+1)} \right|^2 dx \leq \int_{-\infty}^{\infty} \frac{M^2}{x^2+1} dx = M^2\pi.$$

Hence $\frac{F(z)}{z+i} \in H^2(\mathbb{R}^2_+)$. This implies that

$$\lim_{y \searrow 0} \frac{F(x+iy)}{x+i(1+y)}$$

exists a.e., which in turn implies that $\lim F(x+iy)$ exists a.e.

(b) I assume that I can take for granted that $z \mapsto i\frac{1-z}{1+z}$ is a conformal mapping of the unit disc into the upper half plane, since we did this on a previous homework. Then define $G(w) = F(i\frac{1-w}{1+w})$ which is a bounded holomorphic function on D. It now suffices to show that w approaches the unit circle non-tangentially as $y = \text{Re}(z) \to 0$, where w is now given by the inverse mapping

$$w = \frac{-x+(1-y)i}{x+(1+y)i}.$$

Then

$$|w|^2 = \frac{x^2+(1-y)^2}{x^2+(1+y)^2}$$

by straightforward arithmetic. Now if w were approaching the unit circle in a tangential manner, we would have $\frac{d|w|^2}{dy}|_{y=0} = 0$. However,

$$\frac{d|w|^2}{dy} = \frac{4(y^2-x^2-1)}{(x^2+(1-y)^2}$$

which is nonzero at $y = 0$.

□

Exercise 4: Consider $F(z) = e^{i/z}/(z+i)$ in the upper half-plane. Note that $F(x+iy) \in L^2(\mathbb{R})$, for each $y > 0$ and $y = 0$. Observe also that $F(z) \to 0$ as $|z| \to \infty$. However, $F \notin H^2(\mathbb{R}^2_+)$. Why?

Solution. For any fixed $y > 0$,

$$\left| \frac{e^{i/(x+iy)}}{x+i(1+y)} \right|^2 \leq \frac{e^{1/y}}{x^2+1}$$

which is integrable, so $F(x+iy) \in L^2(\mathbb{R})$. For $y = 0$,

$$\left| \frac{e^{i/x}}{x+i} \right|^2 = \frac{1}{|x+i|^2} = \frac{1}{x^2+1}$$

which is again integrable. Also, as $|z| \to \infty$, the numerator approaches 1 in magnitude while the denominator becomes infinite. However, $F \notin H^2$, as is suggested by the fact that our bound includes an $e^{1/y}$ term, which blows up. The problem, of course, is that F is not bounded in the upper half plane; it has an essential singularity at 0, and Picard's theorem tells us that it takes on every complex value (except possibly 1) in every neighborhood of the origin. □

Exercise 6: Suppose Ω is an open set in $\mathbb{C} = \mathbb{R}^2$, and let $\mathcal{H}$ be the subspace of $L^2(\Omega)$ consisting of holomorphic functions on Ω. Show that $\mathcal{H}$ is a closed subspace of $L^2(\Omega)$, and hence is a Hilbert space with inner product

$$(f, g) = \int_\Omega f(z)\bar{g}(z)dxdy, \quad \text{where } z = x + iy.$$

Solution. For any $f \in \mathcal{H}$, $f^2 \in \mathcal{H}$ as well, since the square of an analytic function is analytic. Now by the mean value property, for any $z \in \Omega$ and $r \leq d(z, \Omega^c)$,

$$f(z)^2 = \frac{1}{\pi r^2} \iint\limits_{|\zeta - z| \leq r} f(\zeta)^2 dA,$$

whence

$$|f(z)|^2 \leq \frac{1}{\pi r^2} \iint\limits_{|\zeta - z| \leq r} |f(\zeta)|^2 dA \leq \frac{1}{\pi r^2}\|f\|^2.$$

Now on any compact $K \subset \Omega$, there is a minimum value $r_0 > 0$ of $d(z, \Omega^c)$ for $z \in K$. (This is because the distance between a compact set and a closed set always attains a nonzero minimum.) Then we have $|f(z)| \leq \frac{\sqrt{\pi}}{r_0}\|f\|$ for all $z \in K$ and $f \in L^2(\Omega)$. So if $\{f_n\}$ is a Cauchy sequence in $L^2(\Omega)$, then $\|f_m - f_n\| \to 0$, whence $|f_m - f_n| \to 0$ as well. Thus, $\{f_n\}$ converges uniformly on any compact subset of Ω. Now it is a theorem in complex analysis that the uniform limit of analytic functions is analytic; this may be proved, for example, by using the ML estimate to show that the integral of the limit around any contour is zero, and then applying Morera's theorem. (See e.g. Gamelin p. 136.) This theorem works on any domain, e.g. the interior of any compact disc contained in Ω. This allows us to prove that the limit of $\{f_n\}$ is analytic at each point in Ω, so it's analytic on Ω. $\square$

Exercise 7: Following up on the previous exercise, prove:

(a) If $\{\phi_n\}_{n=0}^\infty$ is an orthonormal basis of $\mathcal{H}$, then

$$\sum_{n=0}^\infty |\phi_n(z)|^2 \leq \frac{c^2}{d(z, \Omega^c)}^2 \quad \text{for } z \in \Omega.$$

(b) The sum

$$B(z, w) = \sum_{n=0}^\infty \phi_n(z)\bar{\phi}_n(w)$$

converges absolutely for $(z, w) \in \Omega \times \Omega$, and is independent of the choice of the orthonormal basis $\{\phi_n\}$ of $\mathcal{H}$.

(c) To prove (b) it is useful to characterize the function $B(z, w)$, called the **Bergman kernel**, by the following property. Let T be the linear transformation on $L^2(\Omega)$ defined by

$$Tf = \int_\Omega B(z, w)f(w)dudv, \quad w = u + iv.$$

Then T is the orthogonal projection of $L^2(\Omega)$ to $\mathcal{H}$.

(d) Suppose that Ω is the unit disc. Then $f \in \mathcal{H}$ exactly when $f(z) = \sum_{n=0}^{\infty} a_n z^n$, with

$$\sum_{n=0}^{\infty} |a_n|^2 (n+1)^{-1} < \infty.$$

Also, the sequence $\{\frac{z^n(n+1)}{\pi^{1/2}}\}_{n=0}^{\infty}$ is an orthonormal basis of $\mathcal{H}$. Moreover, in this case

$$B(z,w) = \frac{1}{\pi(1-z\bar{w})^2}.$$

Solution.

(a) First we prove a lemma, whose relevance was so kindly pointed out by Prof. Garnett:

Lemma 3. *Let $\{b_n\}_{n=0}^{\infty}$ be a sequence of complex numbers. Then*

$$\sqrt{\sum_{n=0}^{\infty} |b_n|^2} = \sup_{\sum |a_n|^2 \leq 1} \left| \sum_{n=0}^{\infty} a_n b_n \right|.$$

Proof. If $\sum |b_n|^2 < \infty$ this follows from the Cauchy-Schwarz inequality applied to $\ell^2(\mathbb{N})$, where equality is achieved when $\{a_n\}$ is the unit vector in the same direction (actually, the conjugate) as $\{b_n\}$. Now suppose $\sum |b_n|^2 = \infty$. Then for any N, define the truncated sequence $\tilde{b}^{(N)} = \{\tilde{b}_n^{(N)}\}$ by

$$\tilde{b}_n^{(N)} = \begin{cases} b_n & n \leq N \\ 0 & \text{else.} \end{cases}$$

Then $\tilde{b}^{(N)} \in \ell^2$, so if $a^{(N)} = \{a_n^{(N)}\}_{n=0}^{\infty}$ is the unit vector in the conjugate direction of $\tilde{b}^{(N)}$, we have

$$\left| \sum_{n=0}^{\infty} a_n^{(N)} b_n \right| = \left| \sum_{n=0}^{\infty} a_n^{(N)} \tilde{b}_n^{(N)} \right| = \|\tilde{b}^{(N)}\|.$$

Since this goes to infinity as $N \to \infty$, we have

$$\sup_{\sum \|a_n\|^2 \leq 1} \left| \sum a_n b_n \right| = \infty = \sqrt{\sum_{n=0}^{\infty} |b_n|^2}.$$

$\square$

Returning to the problem at hand, for any sequence a_n with $\sum |a_n|^2 \leq 1$,

$$g(z) = \sum_{n=0}^{\infty} a_n \phi_n(z)$$

is a unit vector in $\mathcal{H}$. Applying problem 6, we have at any fixed $z \in \Omega$ that

$$\left| \sum_{n=0}^{\infty} a_n \phi_n(z) \right| = |g(z)| \leq \frac{\sqrt{\pi}}{d(z, \Omega^c)} \|g\| = \frac{\sqrt{\pi}}{d(z, \Omega^c)}.$$

Applying the lemma with $b_n = \phi_n(z)$, we have

$$\sum_{n=0}^{\infty} |\phi_n(z)|^2 = \left(\sup_{\sum |a_n|^2 \le 1} \left| \sum_{n=0}^{\infty} a_n \phi_n(z) \right| \right)^2 \le \frac{\pi}{d(z, \Omega^c)^2}.$$

(b) The absolute convergence of this sum follows from part (a) and the Cauchy-Schwarz inequality: For fixed values of z and w, $\{|\phi_n(z)|\}$ and $\{\bar{\phi_n}(w)\}$ are vectors in ℓ^2, so by the Cauchy-Schwarz inequality,

$$\sum \phi_n(z) \phi_n\bar{(}w) \le \sqrt{\sum |\phi_n(z)|^2} \sqrt{\sum |\bar{\phi_n}(w)|^2} < \infty.$$

To prove that the sum is independent of the choice of basis, we use part (c). Because integration against this sum is projection onto $\mathcal{H}$, and there is only one projection map, any two such sums must be equal almost everywhere. I'm not 100% sure how to go about showing they are in fact equal everywhere. Certainly $B(z, w)$ is analytic in z and analytic in w (with either variable fixed, it's in $\mathcal{H}$ as a function of the other variable). However, I don't know anything about functions of several complex variables; is a function that's analytic in each variable separately necessarily analytic? Or, more to the point, continuous? Assuming so, continuity plus equality almost everywhere implies equality. Of course, one could say that since $\mathcal{H}$ is being viewed as a subspace of L^2, a.e. equality is all we need for the functions to be the same point in the Hilbert space.

(c) Since $\{\phi_n\}$ is an ONB for the closed subspace $\mathcal{H}$, we can extend it to a basis for all of L^2 by complementing it with another set $\{\psi_k\}$ of orthonormal vectors. For $z \in \Omega$, define $B_z(w) = \overline{B(z, w)}$. Then

$$Tf(z) = \int_{\Omega} B_z(w) f(w) dw = \langle f, B_z \rangle.$$

We can write f in our ONB as $f(w) = \sum_{n=0}^{\infty} a_n \phi_n(w) + \sum_{k=0}^{\infty} b_k \psi_k(w)$ whence

$$\begin{aligned} Tf(z) &= \langle f, B_z \rangle \\ &= \left\langle \left(\sum_{n=0}^{\infty} a_n \phi_n + \sum_{k=0}^{\infty} b_k \psi_k \right), \left(\sum_{j=0}^{\infty} \bar{\phi_j}(z) \phi_j \right) \right\rangle \\ &= \sum_{j,n} a_n \overline{\phi_j(z)} \langle \phi_n, \phi_j \rangle + \sum_{k,j} b_k \overline{\phi_j(z)} \langle \phi_k, \phi_j \rangle \\ &= \sum_n a_n \overline{\phi_n(z)}. \end{aligned}$$

This is the formula for projection onto a closed subspace–T erases all the components in the orthogonal complement.

(d) The set $\{\phi_n\} = \left\{ z^n \sqrt{\frac{n+2}{\pi}} \right\}$ is orthonormal since

$$\langle \phi_n, \phi_n \rangle = \frac{n+1}{\pi} \int_D |z^2 n| dA = \frac{n+1}{\pi} \int_{r=0}^{1} \int_{\theta=0}^{2\pi} r^2 n r dr d\theta = 2(n+1) \frac{r^{2n+2}}{2n+2} \Big|_0^1 = 1$$

and

$$\langle \phi_n, \phi_\rangle = \frac{\sqrt{(m+1)(n+1)}}{\pi} \int_{r=0}^{1} \int_{\theta=0}^{2\pi} r^{n+m} e^{2\pi i(n-m)} r dr d\theta = 0$$

for $m \neq n$. Now since every analytic function has a power series expansion, any analytic function can be written as $\sum b_n \phi_n$. This proves that $\{\phi_n\}$ is a basis for $\mathcal{H}$, and also gives us the condition for an analytic function to be in L^2: $\sum |b_n|^2 < \infty \Leftrightarrow \sum \frac{|a_n|^2}{n+1} < \infty$, since $b_n = a_n \sqrt{\frac{\pi}{n+1}}$.

To obtain an expression for $B(z, w)$, we first note that for any complex number ζ with $|\zeta| < 1$,

$$\frac{1}{(1-\zeta)^2} = \sum_{n=0}^{\infty} (n+1)\zeta^n.$$

This may be obtained by differentiating the series $\frac{1}{\zeta} = \sum \zeta^n$ termwise, or by squaring it and collecting like terms. Both are justified by the uniform absolute convergence of this series on compact subdisks of the unit disk. Then

$$B(z, w) = \sum_{n=0}^{\infty} \sqrt{\frac{n+1}{\pi}} z^n \sqrt{\frac{n+1}{\pi}} \bar{w}^n = \frac{1}{\pi} \sum_{n=0}^{\infty} (n+1)(z\bar{w})^n = \frac{\pi}{(1-z\bar{w})^2}.$$

□

Exercise 8: Continuing with Exercise 6, suppose Ω is the upper half-plane $\mathbb{R}^2_+$. Then every $f \in \mathcal{H}$ has the representation

$$f(z) = \sqrt{4\pi} \int_0^{\infty} \hat{f}_0(\xi) e^{2\pi i \xi z} d\xi, \quad z \in \mathbb{R}^2_+,$$

where $\int_0^\infty |\hat{f}_0(\xi)|^2 \frac{d\xi}{\xi} < \infty$. Moreover, the mapping $\hat{f}_0 \to f$ given by this formula is a unitary mapping from $L^2((0,\infty), \frac{d\xi}{\xi})$ to $\mathcal{H}$.

Solution. Following the proof of Theorem 2.1 on page 214, we define $\hat{f}_y(\xi)$ to be the Fourier transform of the L^2 function $f(x+iy)$. (We know $f(x+iy)$ is an L^2 function of x for almost all y since

$$\|f\|_2^2 = \int \left(\int |f(x+iy)|^2 dx \right) dy$$

so that $\int |f(x+iy)|^2 dx$ is an integrable function of y, and therefore finite almost everywhere. Then we can show that $\hat{f}_y(\xi) e^{2\pi y \xi}$ is independent of y using exactly the same proof in the book. (Our proof of the boundedness of f on closed half-planes changes slightly: we now have

$$|f(\zeta)|^2 = \frac{1}{\delta^2} \int_{|z|<\delta} |f(\zeta+z)|^2 dx dy \leq \frac{1}{\delta^2} \|f\|_2^2.$$

Other than that the proof requires no modification.) Having established this, we can then define $\hat{f}_0(\xi)$ to be the function that equals $\hat{f}_y(\xi) e^{2\pi y \xi}$ for

almost all y. The Plancherel formula then gives us

$$\int_{-\infty}^{\infty} |f(x+iy)|^2 dx = \int_{-\infty}^{\infty} |\hat{f}_0(\xi)|^2 e^{-4\pi y\xi} d\xi.$$

This tells us that $\hat{f}_0(\xi) = 0$ for a.a. $\xi < 0$ (since the integral in ξ is infinite for $\xi < 0$), and also gives us the relation

$$\begin{aligned}
\|f\|_2^2 &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |f(x+iy)|^2 dxdy \\
&= \int_{-\infty}^{\infty}\int_{0}^{\infty} |\hat{f}_0(\xi)|^2 e^{-4\pi y\xi} d\xi dy \\
&\overset{\text{Tonelli}}{=} \int_{0}^{\infty} |\hat{f}_0(\xi)|^2 \int_{-\infty}^{\infty} e^{-4\pi y\xi} dy d\xi \\
&= \int_{0}^{\infty} |\hat{f}_0(\xi)|^2 \frac{1}{4\pi\xi} d\xi
\end{aligned}$$

This tells us that $\|f\| = \|\frac{1}{\sqrt{4\pi}} f_0\|_{L^2((0,\infty),d\xi/\xi)}$. We also have by Fourier inversion that $f(z) = \sqrt{4\pi}\int_0^\infty \frac{1}{\sqrt{4\pi}}\hat{f}_0(\xi)e^{2\pi iz}d\xi$. If we replace $\hat{f}_0$ by $\frac{1}{\sqrt{4\pi}}\hat{f}_0$, we will have a unitary map $\hat{f}_0 \to f$, and

$$f(z) = \sqrt{4\pi}\int_0^\infty \hat{f}_0(\xi)e^{2\pi i\xi z}d\xi.$$

□

Exercise 9: Let H be the Hilbert transform. Verify that

(a) $H^* = -H$, $H^2 = -I$, and H is unitary.
(b) If τ_h denotes the translation operator, $\tau_h(f)(x) = f(x-h)$, then H commutes with τ_h, $\tau_h H = H\tau_h$.
(c) if δ_a denotes the dilation operator, $\delta_a(f)(x) = f(ax)$ with $a > 0$, then H commutes with δ_a, $\delta_a H = H\delta_a$.

Solution.

(a) Since the projection P and the identity I are both self-adjoint, $2P - I$ is self-adjoint, so $H = -i(2P-I)$ is skew-adjoint.
(b) Since H is a linear combination of I and P, it suffices to verify that both of these commute with τ_h. For I this is trivial. For P, we have

$$\widehat{P(\tau_h f)}(\xi) = \chi(\xi)\widehat{\tau_h f}(\xi) = \chi(\xi)e^{2\pi ih}\hat{f}(\xi)$$

and

$$\widehat{\tau_h P(f)}(\xi) = e^{2\pi ih}\widehat{Pf}(\xi) = e^{2\pi ih}\chi(\xi)\hat{f}(\xi).$$

Since the Fourier transform on L^2 is invertible, this implies $\tau_h Pf = P\tau_h f$, so P commutes with τ_h.
(c) Again, it suffices to verify that P commutes with dilations.

$$\widehat{\delta_a P(f)}(\xi) = a\widehat{Pf}(a\xi) = a\chi(a\xi)\hat{f}(a\xi) = a\chi(\xi)\hat{f}(a\xi)$$

where $\chi(a\xi) = \chi(\xi)$ because $a > 0$. Similarly,

$$\widehat{P\delta_a f}(\xi) = \chi(\xi)\widehat{\delta_a f}(\xi) = \chi(\xi)a\hat{f}(a\xi).$$

Hence P commutes with dilations.

□

Exercise 15: Suppose $f \in L^2(\mathbb{R}^d)$. Prove that there exists $g \in L^2(\mathbb{R}^d)$ such that

$$\left(\frac{\partial}{\partial x}\right)^\alpha f(x) = g(x)$$

in the weak sense, if and only if

$$(2\pi i\xi)^\alpha \hat{f}(\xi) = \hat{g}(\xi) \in L^2(\mathbb{R}^d).$$

Solution. (Help from Kenny Maples.) Let $L = \left(\frac{\partial}{\partial x}\right)^\alpha$. Then $L^* = (-1)^{|\alpha|}\left(\frac{\partial}{\partial x}\right)^\alpha$. Note in particular that

$$\overline{\widehat{L^*\psi}(\xi)} = \overline{(-1)^{|\alpha|}\widehat{\left(\frac{\partial}{\partial x}\right)^\alpha \psi}(\xi)} = \overline{(-1)^{|\alpha|}(2\pi i\xi)^\alpha\hat{\psi}(\xi)} = (2\pi i\xi)^\alpha\overline{\hat{\psi}(\xi)}.$$

Now suppose $\hat{g} = \hat{f}(\xi)(2\pi i\xi)^\alpha \in L^2$. Define $g \in L^2$ as the inverse Fourier transform of $\hat{g}$. Using Plancherel's identity, for any $\psi \in C_0^\infty$ we have

$$\begin{aligned}
\langle g, \psi\rangle &= \langle \hat{g}, \hat{\psi}\rangle \\
&= \int \hat{g}(\xi)\overline{\hat{\psi}(\xi)}d\xi \\
&= \int \hat{f}(\xi)(2\pi i\xi)^\alpha\overline{\hat{\psi}(\xi)}d\xi \\
&= \int \hat{f}(\xi)\overline{\widehat{L^*\psi}(\xi)}d\xi \\
&= \langle \hat{f}, \widehat{L^*\psi}\rangle \\
&= \langle f, L^*\psi\rangle.
\end{aligned}$$

Hence $g = Lf$ weakly.

Conversely, suppose there exists $g \in L^2$ such that $g = Lf$ weakly. Using Plancherel again,

$$\begin{aligned}
\int \hat{g}(\xi)\overline{\hat{\psi}(\xi)}d\xi &= \langle \hat{g}, \hat{\psi}\rangle \\
&= \langle g, \psi\rangle \\
&= \langle f, L^*\psi\rangle \\
&= \langle \hat{f}, \widehat{L^*\psi}\rangle \\
&= \int \hat{f}(\xi)(2\pi i\xi)^\alpha\overline{\hat{\psi}(\xi)}d\xi.
\end{aligned}$$

Since this is true for all $\psi \in C_0^\infty$, we must have $\hat{g}(\xi) = \hat{f}(\xi)(2\pi i\xi)^\alpha$ a.e. Since $g \in L^2$, $\hat{g} \in L^2$ by Plancherel, so $\hat{f}(\xi)(2\pi i\xi)^\alpha = \hat{g}(\xi) \in L^2$. □

Chapter 5.6, Page 259

Problem 6: This problem provides an example of the contrast between analysis on $L^1(\mathbb{R}^d)$ and $L^2(\mathbb{R}^d)$.

Recall that if f is locally integrable on $\mathbb{R}^d$, the maximal function f^* is defined by

$$f^*(x) = \sup_{x\in B} \frac{1}{m(B)} \int_B |f(y)|dx,$$

where the supremum is taken over all balls containing the point x.

Complete the following outline to prove that there exists a constant C so that

$$\|f^*\|_2 \le C\|f\|_2.$$

In other words, the map that takes f to f^* (although not linear) is bounded on $L^2(\mathbb{R}^d)$. This differs notably from the situation in $L^1(\mathbb{R}^d)$, as we observed in Chapter 3.

(a) For each $\alpha > 0$, prove that if $f \in L^2(\mathbb{R}^d)$, then

$$m(\{x : f^*(x) > \alpha\}) \le \frac{2A}{\alpha} \int_{|f|>\alpha/2} |f(x)|dx.$$

Here, $A = 3^d$ will do.

(b) Show that

$$\int_{\mathbb{R}^d} |f^*(x)|^2 dx = 2\int_0^\infty \alpha m(E_\alpha) d\alpha,$$

where $E_\alpha = \{x : f^*(x) > \alpha\}$.

(c) Prove that $\|f^*\|_2 \le C\|f\|_2$.

Solution.

(a) Let $G_\alpha = \{x : |f(x)| > \frac{\alpha}{2}\}$. Then $1 \le \frac{2}{\alpha}|f|$ on G_α, so

$$\int_{G_\alpha} |f(y)|dy \le \int_{G_\alpha} \frac{2}{\alpha}|f(y)|^2 dy \le \frac{2}{\alpha}\|f\|_2 < \infty.$$

Now let $E_\alpha = \{x : f^*(x) > \alpha\}$. For any $x \in E_\alpha$, $\exists B_x$ with $x \in B_x$ and

$$m(B_x) < \frac{1}{\alpha}\int_{B_x} |f(y)|dy < \frac{1}{\alpha}\left(\frac{\alpha}{2}m(B_x) + \int_{G_\alpha\cap B_x} |f(y)|dy\right) \Rightarrow m(B_x) < \frac{2}{\alpha}\int_{G_\alpha\cap B_x} |f(y)|dy.$$

Here we have broken up the integral into the integral over the portion of B_x where $|f| \le \frac{\alpha}{2}$ and the region where $|f| > \frac{\alpha}{2}$ and bounded each portion. Now let K be any compact subset of E_α; then K is covered by finitely many balls $B_{x_1}, \dots, B_{x_N}$. By the Covering Lemma, there exists a subcollection $B_{x_{n_1}}, \dots, B_{x_{n_M}}$ such that

$$m\left(\bigcup_{i=1}^N B_i\right) \le 3^d \sum_{j=1}^M m(B_{x_{n_j}}).$$

Then

$$m(K) \le 3^d \sum_{j=1}^M m(B_{x_{n_j}}) \le \frac{2\cdot 3^d}{\alpha}\sum_{j=1}^M \int_{G_\alpha\cap B_{x_{n_j}}} |f(y)|dy \le \frac{2\cdot 3^d}{\alpha}\int_{G_\alpha} |f(y)|dy.$$

By the regularity of Lebesgue measure,

$$m(E_\alpha) = \underset{K\subset E_\alpha \text{ cpct}}{m(K)} \Rightarrow m(E_\alpha) \le \frac{2\cdot 3^d}{\alpha}\int_{G_\alpha} |f(y)|dy.$$

(b) Using Tonelli's theorem,

$$\int_{\mathbb{R}^d} |f^*(x)|^2 dx = \int_{\mathbb{R}^d} \int_0^\infty \chi_{|f^*(x)|^2 > y} dy dx = \int_0^\infty m\left(\{x : |f^*(x)| > \sqrt{y}\}\right) dy.$$

Substituting $\alpha = \sqrt{y}, dy = 2\alpha d\alpha$, this equals

$$2 \int_0^\infty \alpha m(\{|f^*(x)| > \alpha\}) d\alpha.$$

(c)

□

Chapter 6.7, Page 312

Exercise 3: Consider the exterior Lebesgue measure m_* introduced in Chapter 1. Prove that a set $E \subset \mathbb{R}^d$ is Carathéodory measurable if and only if E is Lebesgue measurable in the sense of Chapter 1.

Exercise 4: Let r be a rotation of $\mathbb{R}^d$. Using the fact that the mapping $x \mapsto r(x)$ preserves Lebesgue measure (see Problem 4 in Chapter 2 and Exercise 26 in Chapter 3), show that it induces a measure-preserving mape of the sphere S^{d-1} with its measure $d\sigma$.

Solution. Let $E \subset S^{d-1}$. By definition, $\sigma(E) = dm(\tilde{E})$ where $\tilde{E}$ is the union of all radii with endpoints in E. Then if r is a rotation of $\mathbb{R}^d$, $\sigma(rE) = dm(\tilde{rE})$ by definition. But $\tilde{rE} = r\tilde{E}$ since

$$\begin{aligned} x \in \tilde{rE} &\Leftrightarrow x = \rho\theta \quad \text{for some} \rho \leq 1, \theta \in rE \\ &\Leftrightarrow x = \rho r(\alpha), \rho \leq 1, \alpha \in E \\ &\Leftrightarrow x = r(\rho\alpha) \text{ (since rotations are linear)} \\ &\Leftrightarrow x \in r(\tilde{E}) \end{aligned}$$

Thus, $m(rE) = dm(r\tilde{E}) = dm(\tilde{E}) = m(E)$, so r preserves measures on the sphere. □

Exercise 5: Use the polar coordinate formula to prove the following:

(a) $\int_{\mathbb{R}^d} e^{-\pi|x|^2} dx = 1$, when $d = 2$. Deduce from this that the same identity holds for all d.

(b) $\left(\int_0^\infty e^{-\pi r^2} r^{d-1} dr\right) \sigma(S^{d-1}) = 1$, and as a result, $\sigma(S^{d-1}) = 2\pi^{d/2}/\Gamma(d/2)$.

(c) If B is the unit ball, $v_d = m(B) = \pi^{d/2}/\Gamma(d/2+1)$, since this quantity equals $\left(\int_0^1 r^{d-1} dr\right) \sigma(S^{d-1})$.

Solution.

(a) For $d = 2$, we have by polar coordinates

$$\int_{\mathbb{R}^d} e^{-\pi|x|^2} dx = \int_{S^1} \int_0^\infty e^{-\pi r^2} r dr d\theta = 2\pi \int_0^\infty e^{-\pi r^2} r dr = -e^{-\pi r^2}|_0^\infty = 1.$$

Note that for general d,

$$\int_{\mathbb{R}^d} e^{-\pi|x|^d} dx = \int_{\mathbb{R}^d} e^{-\pi(x_1^d + \cdots + x_d^d)} dx = \int \ldots \int e^{-\pi x_1^d} \ldots e^{-\pi x_d^d} dx_1 \ldots dx_d = \left(\int_{\mathbb{R}^1} e^{-\pi x^2} dx\right)^d$$

by Tonelli's Theorem. Since we have calculated that this equals 1 for $d = 2$, it follows that $\int_{\mathbb{R}^1} e^{-\pi x^2} dx = 1$, whence the integral over $\mathbb{R}^d$ is 1 for all d.

(b) Using integration by parts, for $d \geq 3$,

$$\int_0^\infty e^{-\pi r^2} r^{d-1} dr = -\frac{e^{-\pi r^2}}{2\pi} r^{d-2}\Big|_0^\infty + \int_0^\infty \frac{e^{-\pi r^2}}{2\pi}(d-2)r^{d-3} dr.$$

Since $\sigma(S^{d-1})$ is just the reciprocal of this first integral (which follows immediately from applying the polar coordinates formula to the result in part (a)), it follows that $\sigma(S^{d-1}) = \frac{2\pi}{d-2}\sigma(S^{d-3})$. We now prove the formula $\sigma(S^{d-1}) = 2\pi^{d/2}/\Gamma(d/2)$ by induction. The base cases are $\sigma(S^1) = 2\pi = 2\pi^{2/2}/\Gamma(2/2)$ and $\sigma(S^2) = 4\pi = 2\pi^{3/2}/(\sqrt{\pi}/2) = 2\pi^{3/2}/\Gamma(3/2)$ since $\Gamma(3/2) = 1/2\Gamma(1/2) = 1/2\sqrt{\pi}$. Now if we let $a_d = 2\pi^{d/2}/\Gamma(d/2)$, then $a_d = \frac{2\pi}{d-2}\pi^{(d-2)/2}/((d-2)/2\Gamma((d-2)/2) = \frac{2\pi}{d-2}a_{d-2}$. Since a_d and $\sigma(S^{d-1})$ satisfy the same recurrence and initial conditions, they are equal for all d.

(c) By polar coordinates,

$$m(B) = \sigma(S^{d-1})\int_0^1 r^{d-1} dr = 2\pi^{d/2}/\Gamma(d/2)\frac{1}{d} = \frac{2\pi^{d-2}}{2\frac{d}{2}\Gamma(d/2)} = \frac{\pi^{d/2}}{\Gamma(d/2+1)}.$$

□

Exercise 8: The fact that the Lebesgue measure is uniquely characterized by its translation invariance can be made precise by the following assertion: If μ is a Borel measure on $\mathbb{R}^d$ that is translation-invariant, and is finite on compact sets, then μ is a multiple of Lebesgue measure m. Prove this theorem by proceeding as follows.

(a) Suppose Q_a denotes a translate of the cube $\{x : 0 < x_j \leq a, j = 1, \ldots, d\}$ of side length a. If we let $\mu(Q_1) = c$, then $\mu(Q_{1/n}) = cn^{-d}$ for each integer n.

(b) As a result μ is absolutely continuous with respect to m, and there is a locally integrable function f such that

$$\mu(E) = \int_E f dx.$$

(c) By the differentiation theorem (Corollary 1.7 in Chapter 3) it follows that $f(x) = c$ a.e., and hence $\mu = cm$.

Solution.

(a) Because Q_1 is a disjoint union of n^d translates of the cube $Q_{1/n}$, $\mu(Q_1) = n^d\mu(Q_{1/n}) \Rightarrow \mu(Q_{1/n}) = n^{-d}\mu(Q_1) = cn^{-d}$.

(b) Let E be Borel measurable with $m(E) = 0$. Then for any $\epsilon > 0$ there is an open set U with $m(U \setminus E) < \epsilon \Rightarrow m(U) < \epsilon$. We can write U as a countable disjoint union of cubes Q_j whose side lengths are of the form $1/n$, for example by decomposing U into dyadic rational cubes as described on pp. 7-8. Then $\mu(Q_j) = cm(Q_j)$ by part (a), so $\mu(U) = \sum \mu(Q_j) = \sum cm(Q_j) = c\sum m(Q_j) = cm(U) < c\epsilon$. This can be done for any ϵ, so $\mu(E) = 0$. Thus, μ is absolutely continuous

wrt m, so there exists a locally integrable Borel measurable function f such that $\mu(E) = \int_E f dm$.

(c) Let x be a Lebesgue point of f. Let Q_n be a series of dyadic rational cubes containing x. (Just to be clear, these are "half-open" dyadic rational cubes, i.e. ones of the form $\frac{m_i}{2^{n_i}} \leq x_i < \frac{m_i+1}{2^{n_i}}$ for $i = 1, \ldots, d$.) Then Q_n shrinks regularly to x because the ratio of a cube to the circumscribing ball is constant. For each cube we have

$$\frac{1}{m(Q_n)} \int_{Q_n} f d\mu = \frac{1}{m(Q_n)} \mu(Q_n) = \frac{1}{m(Q_n)} cm(Q_n) = c,$$

so $f(x) = c$ at every Lebesgue point x of f (hence a.e.).

□

Exercise 9: Let $C([a, b])$ denote the vector space of continuous functions on the closed and bounded interval $[a, b]$. Suppose we are given a Borel measure μ on this interval, with $\mu([a, b]) < \infty$. Then

$$f \mapsto \ell(f) = \int_a^b f(x) d\mu(x)$$

is a linear functional on $C([a, b])$, with ℓ positive in the sense that $\ell(f) \geq 0$ if $f \geq 0$.

Prove that, conversely, for any linear functional ℓ on $C([a, b])$ that is positive in the above sense, there is a unique finite Borel measure μ so that $\ell(f) = \int_a^b f d\mu$ for $f \in C([a, b])$.

Solution. Define the notation $f \prec u$ to mean $0 \leq f \leq 1$ and $f = 1$ on $[a, u]$. Let

$$F(u) = \underset{f \prec u}{\ell}(f).$$

Then F is increasing on $[a, b]$, because for $u' > u$, $f \prec u' \Rightarrow f \prec u$ so $F(u')$ is the infimum of a smaller class of sets. To show F is right continuous, it suffices to show that for every $f \prec u$ and $\epsilon > 0$ there exists a $u' > u$ and $f' \prec u'$ with $\ell(f') < \ell(f) + \epsilon$. Let $f \prec u$ and let $C = \ell(f)$. By continuity, the function $(1 + \frac{\epsilon}{C})f$ is greater than 1 in some neighborhood $u < x < x + \delta$. Let $f' \prec u + \frac{\delta}{2}$ and $f'(y) = 0$ for $y \geq \delta$. (Such an f' can be constructed, for example, as piecewise linear, say $f' = 1$ on $[a, u + \delta/2]$, $f'(u + 3\delta/4) = 0$, and f is linear between $\delta/2$ and $3\delta/4$.) Then $f'(x) \leq (1 + \frac{\epsilon}{C})f(x)$ everywhere, so $\ell(f') \leq (1 + \frac{\epsilon}{C})\ell(f) = C + \epsilon$. This proves that F is right continuous. By Theorem 3.5, there exists a unique Borel measure μ on $[a, b]$ such that $\mu((a', b']) = F(b') - F(a')$ for all $a \leq a' < b' \leq b$.

Now we need to show $\ell(f) = \int_a^b f d\mu$ for all $f \in C([a, b])$. Let $L(f) = \int_a^b f d\mu$. Then it suffices to show $\ell(f) \leq L(f)$ since this will imply $-\ell(f) = \ell(-f) \leq L(-f) = -L(f) \Rightarrow \ell(f) \geq L(f)$. Let $\epsilon > 0$. Because continuous functions can be uniformly approximated by step functions, we may choose a step function $f_\epsilon \geq f$ with $L(f_\epsilon) < L(f) + \epsilon$. Write

$$f_\epsilon = \sum c_k \chi_{(a_k, b_k]}.$$

WLOG we may assume that the intervals $(a_k, b_k]$ are disjoint. Choose $a'_k > a_k$ and let $f'_\epsilon = \sum c_k \chi_{(a'_k, b_k]}$. Now by the definition of F there exist continuous g_k and h_k with $g_k \prec b_k$, $h_k \prec a'_k$, and $\ell(h_k) - F(a'_k) <$

$\ell(g_k) - F(b_k) < \frac{\epsilon}{2^k}$. WLOG we may also assume $f < c_k(1-h_k)$ on (a_k, a'_k) since otherwise we may take

$$c_k(1-h'_k) = \begin{cases} max(f, c_k(1-h_k)) & a_k < x < a'_k \\ c_k & a'_k \leq x < b_k \end{cases}$$

and the function h'_k defined by these relations will also be continuous, satisfy $h'_k \prec a'_k$, and have $h'_k < h_k \Rightarrow \ell(h'_k) < \ell(h_k)$. Now let

$$\tilde{f}_\epsilon = \sum c_k(g_k - h_k).$$

Then we have $\tilde{f}_\epsilon \geq f$ by the above remarks concerning h_k. Note also that

$$\ell(\tilde{f}_\epsilon) = \sum c_k(\ell(g_k) - \ell(h_k)) < \sum c_k(F(b_k) - F(a'_k)) + \epsilon = L(f'_\epsilon) + \epsilon.$$

Since we also have the relations $\tilde{f}_\epsilon \geq f$ and $f_\epsilon \geq f'_\epsilon$, and both ℓ and L are positive,

$$\ell(f) < \ell(\tilde{f}_\epsilon) < L(f'_\epsilon) + \epsilon < L(f_\epsilon) + \epsilon < L(f) + 2\epsilon.$$

This is true for all ϵ, so $\ell(f) \leq L(f)$. □

Exercise 10: Suppose ν, ν_1, ν_2 are signed measures on $(X, \mathcal{M})$ and μ a (positive) measure on $\mathcal{M}$. Using the symbols $\perp$ and $\ll$ defined in Section 4.2, prove:

(a) If $\nu_1 \perp \mu$ and $\nu_2 \perp \mu$, then $\nu_1 + \nu_2 \perp \mu$.
(b) If $\nu_1 \ll \mu$ and $\nu_2 \ll \mu$, then $\nu_1 + \nu_2 \ll \mu$.
(c) $\nu_1 \perp \nu_2$ implies $|\nu_1| \perp |\nu_2|$.
(d) $\nu \ll |\nu|$.
(e) If $\nu \perp \mu$ and $\nu \ll \mu$, then $\nu = 0$.

Solution.

(a) Let disjoint A_1 and B_1 be chosen such that $\nu_1(E) = \nu_1(A_1 \cap E)$ and $\mu(E) = \mu(B_1 \cap E)$ for all measurable E. Similarly, choose A_2 and B_2 disjoint with $\nu_2(E) = \nu_2(A_2 \cap E)$ and $\mu(E) = \mu(B_2 \cap E)$. Let $A = A_1 \cup A_2$ and $B = B_1 \cap B_2$. Note that A and B are disjoint because $A_1 \cap B \subset A_1 \cap B_1 = \emptyset$ and similarly for A_2. Then for any measurable E, $\mu(E) = \mu(E \cap B_1) = \mu(E \cap B) + \mu(E \cap (B_1 \setminus B_2)$. But $\mu(B_1 \setminus B_2) = \mu((B_1 \setminus B_2) \cap B_2) = 0$, so $\mu(E) = \mu(E \cap B)$. Similarly, $\nu_1(E) = \nu_1(E \cap A_1) = \nu_1(E \cap A) - \nu_1(E \cap (A \setminus A_1))$, but $\nu_1(A \setminus A_1) = \nu_1((A \setminus A_1) \cap A_1) = 0$, so $\nu_1(E) = \nu_1(E \cap A)$ and by the same token, $\nu_2(E) = \nu_2(E \cap A)$, so $(\nu_1 + \nu_2)(E) = (\nu_1 + \nu_2)(E \cap A)$. Thus, μ and $\nu_1 + \nu_2$ are supported on the disjoint sets A and B.
(b) $\mu(E) = 0 \Rightarrow \nu_1(E) = \nu_2(E) = 0 \Rightarrow (\nu_1 + \nu_2)(E) = 0$.
(c) Choose disjoint A and B such that $\nu_1(E) = \nu_1(E \cap A)$ and $\nu_2(E) = \nu_2(E \cap B)$ for all measurable E. Then

$$|\nu_1|(E) = \sup \sum_j |\nu_1(E_j)| = \sup \sum_j |\nu_1(E_j \cap A)| = |\nu_1|(E \cap A)|$$

and similarly $|\nu_2|(E) = |\nu_2|(E \cap B)$. Hence $|\nu_1|$ and $|\nu_2|$ are supported on the disjoint sets A and B.
(d) $|\nu|(E) = 0 \Rightarrow \sup \sum |\nu(E_j)| = 0 \Rightarrow \nu(E_j) = 0$ for all subsets $E_j \subset E \Rightarrow \nu(E) = 0$.

(e) Let disjoint A and B be chosen with $\nu(E) = \nu(E \cap A)$ and $\mu(E) = \mu(E \cap B)$. Then for any measurable E, $\mu(E \cap A) = \mu((E \cap A) \cap B) = 0$ because A and B are disjoint. Then $\nu(E) = \nu(E \cap A) = 0$ because $\mu(E \cap A) = 0$ and $\nu \ll \mu$.

□

Exercise 11: Suppose that F is an increasing normalized function on $\mathbb{R}$, and let $F = F_A + F_C + F_J$ be the decomposition of F in Exercise 24 of Chapter 3; here F_A is absolutely continuous, F_C is continuous with $F_C' = 0$ a.e., and F_J is a pure jump function. Let $\mu = \mu_A + \mu_C + \mu_J$ with μ, μ_A, μ_C, and μ_J the Borel measures associated to F, F_A, F_C, and F_J respectively. Verify that:

(i) μ_A is absolutely continuous with respect to Lebesgue measure and $\mu_A(E) = \int_E F'(x)dx$ for every Lebesgue measurable set E.

(ii) As a result, if F is absolutely continuous, then $\int f d\mu = \int f dF = \int f(x)F'(x)dx$ whenever f and fF' are integrable.

(iii) $\mu_C + \mu_J$ and Lebesgue measure are mutually singular.

Solution.

(i) By definition

$$\begin{aligned}
\mu_A(E) &= \inf_{E \subset \cup(a_j,b_j]} \sum F_A(b_j) - F_A(a_j) \\
&= \inf_{E \subset \cup(a_j,b_j]} \sum \int_{a_j}^{b_j} F'(x)dx \\
&\geq \inf_{E \subset \cup(a_j,b_j]} \int_{\cup(a_j,b_j]} F'(x)dx \\
&\geq \int_E F'(x)dx.
\end{aligned}$$

To prove the reverse inequality, let $\epsilon > 0$ and use the absolute continuity of the integral to find a $\delta > 0$ such that $m(E) < \delta \Rightarrow \int_E F'(x) < \epsilon$. (In case the assumption that $F' \in L^1$ is a problem, we can treat the intersection of E with each interval $[n, n+1)$ separately.) Now since E is Lebesgue measurable, there is an open set $U \supset E$ such that $m(U \setminus E) < \delta$. Let $\bar{U}$ be constructed by writing U as a disjoint union of open intervals (a_j, b_j) and replacing each with $(a_j, b_j]$. Then $m(\bar{U} \setminus E) = m(U \setminus E)$ and $\int_{\bar{U}\setminus E} F'(x)dx = \int_{U \setminus E} F'(x)dx$ because $\bar{U} \setminus E$ is $U \setminus E$ plus countably many points. Thus

$$\int_{\bar{U}} F'(x)dx = \int_{\bar{U}\setminus E} F'(x)dx + \int_E F'(x)dx \leq \epsilon + \int_E F'(x)dx.$$

But $\bar{U}$ is one of the sets over which the infimum is taken in the definition of $\mu_A(E)$, so $\mu_A(E) \leq \epsilon + \int_E F'(x)dx$. This is true for any ϵ, so $\mu_A(E) = \int_E F'(x)dx$.

(ii) The equation $\int_E f d\mu = \int_E fF'(x)dx$ follows immediately from (a) in the case where f is a characteristic function. By the linearity of the integral, it holds for f a simple function as well. The result for non-negative f follows from the Monotone Convergence Theorem: Choose

simple $f_n \nearrow f$. Then

$$\int_E f d\mu = \int_E (\lim f_n) d\mu = \lim \int_E f_n d\mu = \lim \int_E f_n(x)F'(x)dx = \int_E (\lim f_n)(x)F'(x)dx = \int_E f(x)F'(x)dx.$$

Finally, linearity allows us to extend to functions in $L^1(\mu)$. Note that $f \in L^1(\mu)$ iff $fF' \in L^1(dx)$ because both $\int |f| d\mu$ and $\int |fF'| dx$ are defined as suprema of integrals of simple functions, and $\int |c| d\mu = \int |gF'| dx$ for g simple so the suprema are taken over the same sets. The condition that f be integrable is, so far as I can tell, superfluous, unless it means μ-integrable, in which case it's redundant.

(iii) By Exercise 10a, it is sufficient to prove that μ_C and μ_J are both singular wrt Lebesgue measure. Write $F_J(x) = \sum_{k=1}^{\infty} c_k \chi_{[x_k,\infty)}$. Let $A = \{x_k\}$ which is countable and therefore has Lebesgue measure zero, so that Lebesgue measure is supported on A^c. Now A^c is open, so it is covered by countably many intervals $(a_j, b_j]$. (We can write A^c as a countable disjoint union of open intervals, and any open interval is a countable union of half-closed intervals.) Thus, any subset $E \subset A^c$ can be covered by countably many intervals $(a_j, b_j]$ with $F_J(a_j) = F_J(b_j)$, so $\mu_J(E) = 0$. Thus, $\mu_J \perp m$.

The proof that $\mu_C \perp m$ is a bit trickier. (Help from Paul Smith on this part.) We will use the following lemma, taken from page 35 of Folland:

Lemma 4. *If μ_F is the Borel measure corresponding to the increasing, right-continuous function F, then for any μ-measurable set E,*

$$\mu(E) = \inf_{E \subset \cup (a_j, b_j)} \sum \mu((a_j, b_j)).$$

In words, this lemma says that it is equivalent to use coverings of open intervals instead of half-open intervals. This is nice because it enables us to use theorems about open covers. The straightforward but unenlightening proof is in Folland for anyone who cares.

Let $A = \{x \in \mathbb{R} : F'_C(x) = 0\}^c$. Then $m(A) = 0$ by hypothesis, so Lebesgue measure is supported on A^c. We wish to show that $\mu_C(A^c) = 0$, so that μ_C is supported on A. It is sufficient to show that $\mu_C(A^c \cap [0,1]) = 0$ since $A^c = \cup A^c \cap [n, n+1]$ and replacing $F_C(x)$ by $F_C(x-n)$ shifts $A^c \cap [n, n+1]$ to $A^c \cap [0,1]$. Thus, let $B = A^c \cap [0,1]$. Let $\epsilon > 0$. Now for any $x \in B$, $F'_C(x) = 0$, so $\exists h_x > 0$ such that $|F_C(y) - F_C(x)| < \epsilon |y - x|$ for $y \in [x-h, x+h]$. By the Dreiecksungleichung, this implies $F_C(x+h) - F_C(x-h) < 2h\epsilon$. Thus, if we let $I_x = (x-h, x+h)$, we have $\mu(I_x) \le \mu((x-h, x+h]) = F_C(x+h) - F_C(x-h) < 2h\epsilon = \epsilon m(I_x)$. The intervals I_x are an open cover of B, so we can take a countable subcover I_n. (Here we use the fact that every subset of $\mathbb{R}$ is a Lindelöf space. The proof-by-jargon of this fact is that every subspace of a separable metric space is separable, and a metric space is separable iff it is Lindelöf. The direct proof is that every open subset of $\mathbb{R}$ is a countable union of rational intervals, so for each rational interval we can take a member of our open cover (if there is any) which contains that interval, and the resulting countable subcollection will still cover our set.) Next, we shrink the intervals I_n to make $\sum m(I_n) < 3$.

We do this inductively: For a given I_n, if $I_n \subset I_j$ for some $j < n$, we discard I_n entirely; similarly, if $I_j \subset I_n$ for some $j < n$ then we discard I_j. Once this is done, $I_n \cap (\cup_{j<n} I_j)$ is an open subset of I_n, hence a disjoint union of open intervals; but none of these intervals can have both its endpoints within I_n because this would imply $I_j \subset I_n$ for some $j < n$. Hence if $I_n = (a, b)$, then $I_n \cap (\cup_{j<n} I_j) = (a, \lambda) \cup (\gamma, b)$ for some $a \le \lambda < \gamma \le b$. By replacing a with $\max(a, \lambda - \frac{1}{2^{n+1}})$ and b with $\min(b, \gamma + \frac{1}{2^{n+1}})$, we form a new interval $I'_n \subset I_n$ with the property that $m(I_n \cap (\cup_{j<n} I_j)) < \frac{1}{2^n}$. However, $\cup_{j=1}^n I'_j = \cup_{j=1}^n I_j$ because we only delete parts of an interval that are covered by other intervals. We would also like I'_n to still have the property that $\mu(I'_n) < \epsilon m(I'_n)$. Unfortunately, this is only guaranteed as long as I'_n contains the central point x_n from the interval I_n. So far I have not been able to close this hole in the proof.

Overlooking this problem, we have found an open cover $I'_n \supset B$ with the properties that $\mu(I'_n) < \epsilon m(I'_n)$ for each n, and $m(I'_n \cap (\cup_{j<n} I'_j)) < \frac{1}{2^n}$. We may additionally assume that $I'_n \subset [-\frac{1}{2}, \frac{3}{2}]$ for all n since we are only interested in covering $[0, 1]$. This will then imply that $\sum m(I_n) < 3$ because if we write $I'_n = A_n \cup B_n$ where $A_n = I_n \setminus (\cup_{j<n} I'_j)$ and $B_n = I'_n \cap (\cup_{j<n} I'_j)$, then the A_n are disjoint and have union equal to $\cup I'_n$, and

$$\sum m(I'_n) = \sum m(A_n) + m(B_n) = m\left(\cup A_n\right) + \sum m(B_n) \le 2 + \sum \frac{1}{2^n} = 3$$

where $m\left(\cup A_n\right) < 2$ because $\cup A_n \subset [-\frac{1}{2}, \frac{3}{2}]$. This then implies

$$\mu(B) \le \mu\left(\cup I'_n\right) \le \sum \mu(I'_n) \le \epsilon \sum m(I'_n) < 3\epsilon.$$

This is true for any ϵ, so $\mu(B) = 0$.

□

Exercise 14: Suppose $(X_j, \mathcal{M}_j, \mu_j)$, $1 \le j \le k$, is a finite collection of measure spaces. Show that parallel with the case $k = 2$ considered in Section 3 one can construct a product measure $\mu_1 \times \cdots \times \mu_k$ on $X = X_1 \times \cdots \times X_k$. In fact, for any set $E \subset X$ such that $E = E_1 \times \cdots \times E_k$, with $E_j \in \mathcal{M}_j$ for all j, define $\mu_0(E) = \prod_{j=1}^k \mu_j(E_j)$. Verify that μ_0 extends to a premeasure on the algebra $\mathcal{A}$ of finite disjoint unions of such sets, and then apply Theorem 1.5.

Solution. First a hand should at least be waved at the fact that $\mathcal{A}$ is an algebra. It is closed under complements because

$$\begin{aligned}(E_1 \times \cdots \times E_k)^c = &(E_1^c \times X_2 \times \cdots \times X_k) \\ &\cup (E_1 \times E_2^c \times X_3 \times \cdots \times X_k) \\ &\cup \cdots \cup (E_1 \times E_2 \cdots \times E_{k-1} \times E_k^c).\end{aligned}$$

This is a "stopping time" argument: we divide the complement into k sets based on which is the first of the E_k that a point fails to be in. The intersection of two unions of disjoint measurable rectangles is another union

of disjoint measurable rectangles, so we only need to check unions. This follows from

$$\begin{aligned}(E_1 \times \cdots \times E_k) \cup (F_1 \times \cdots \times F_k) = (E_1 \times \cdots \times E_k) \\ \cup (F_1 \setminus E_1 \times F_2 \times \cdots \times F_k) \\ \cup (F_1 \cap E_1 \times F_2 \setminus E_2 \times F_3 \times \cdots \times F_k) \\ \cup (F_1 \cap E_1 \times F_2 \cap E_2 \times F_3 \setminus E_3 \times F_4 \times \cdots \times F_k) \\ \cup \cdots \cup (F_1 \cap E_1 \times \cdots \times F_{k-1} \cap E_{k-1} \times F_k \setminus E_k).\end{aligned}$$

Finally, to show that the extension from rectangles to a premeasure on the algebra $\mathcal{A}$ generated by them is well-defined, let $E_1 \times \cdots \times E_k$ be a measurable rectangle, and suppose $E_1 \times \cdots \times E_k = \cup_j E_1^j \times \cdots \times E_k^j$ where the union is disjoint. This immediately implies

$$\chi_{E_1}(x_1) \ldots \chi_{E_k}(x_k) = \sum_{j=1}^{\infty} \chi_{E_1^j}(x_1) \ldots \chi_{E_k^j}(x_j)$$

for all $(x_1, \ldots, x_k) \in X_1 \times \cdots \times X_k$. We can integrate both sides with respect to x_1, using the monotone convergence theorem to move the integral inside the sum on the RHS, to obtain

$$\mu_1(E_1)\chi_{E_2}(x_2) \ldots \chi_{E_k}(x_k) = \sum_{j=1}^{\infty} \mu_1(E_1^j)\chi_{E_2^j}(x_2) \ldots \chi_{E_k^j}(x_k).$$

We then integrate each side wrt x_2, etc. After doing this k times we obtain

$$\mu_1(E_1) \ldots \mu_k(E_k) = \sum_{j=1}^{\infty} \mu_1(E_1^j) \ldots \mu_k(E_k^j) \Rightarrow \mu_0(E_1 \times \cdots \times E_k) = \sum_{j=1}^{\infty} \mu_0(E_1^j \times \cdots \times E_k^j)$$

as desired.

Since μ_0 is a premeasure on $\mathcal{A}$, it extends to a measure on the σ-algebra generated by $\mathcal{A}$ by Theorem 1.5. □

Exercise 15: The product theory extends to infinitely many factors, under the requisite assumptions. We consider measure spaces $(X_j, \mathcal{M}_j, \mu_j)$ with $\mu_j(X_j) = 1$ for all but finitely many j. Define a **cylinder set** E as

$\{x = (x_j), x_j \in E_j, E_j \in \mathcal{M}_j, E_j = X_j$ for all but finitely many $j\}$.

For such a set define $\mu_0(E) = \prod_{j=1}^{\infty} \mu_j(E_j)$. If $\mathcal{A}$ is the algebra generated by the cylinder sets, μ_0 extends to a premeasure on $\mathcal{A}$, and we can apply Theorem 1.5 again.

Solution. First, note that finite disjoint unions of cylinder sets form an algebra, which is therefore the algebra $\mathcal{A}$. To see this, we can just apply Exercise 14 because of the condition that finitely many indices in the cylinder have $E_j \neq X_j$. For example, to see how unions work, let $\prod E_j$ be a cylinder set (where $E_j = X_j$ for all but finitely many j) and $\prod F_j$ another cylinder set. Then there are finitely many j for which either E_j or F_j is not X_j; we may apply the decomposition from Exercise 14 to these components while leaving the others untouched, and hence obtain a decomposition of $(\prod E_j) \cup (\prod F_j)$ into finitely many disjoint cylinder sets. Similar comments apply to intersections and complements.

To verify that μ_0 extends to a premeasure on $\mathcal{A}$, let $\prod E_j$ be a cylinder set, and suppose

$$\prod_{j=1}^{\infty} E_j = \bigcup_{k=1}^{\infty} \prod_{j=1}^{\infty} E_j^k$$

where the union is disjoint and all but finitely many E_j^k are equal to X_j for any fixed k. The characteristic-function version of this statement is

$$\prod_{j=1}^{\infty} \chi_{E_j}(x_j) = \sum_{k=1}^{\infty} \prod_{j=1}^{\infty} \chi_{E_j^k}(x_j).$$

Integrating both sides with respect to x_1 and using the monotone convergence theorem to move the integral inside the sum on the right,

$$\mu_1(E_1) \prod_{j=2}^{\infty} \chi_{E_j}(x_j) = \sum_{k=1}^{\infty} \mu_1(E_1^k) \prod_{j=2}^{\infty} \chi_{E_j^k}(x_j).$$

Repeating the process ℓ times, we have

$$\mu_1(E_1) \dots \mu_\ell(E_\ell) \prod_{j=\ell+1}^{\infty} \chi_{E_j}(x_j) = \sum_{k=1}^{\infty} \mu_1(E_1^k) \dots \mu_\ell(E_\ell^k) \prod_{j=\ell+1}^{\infty} \chi_{E_j^k}(x_j).$$

As $\ell \to \infty$, the LHS approaches $\mu_0(\prod E_j)$; in fact, it will equal it after a finite number of steps because all but finitely many E_j equal X_j and $\mu_j(X_j) = 1$ for all but finitely many X_j. For the RHS, we apply monotone convergence again (this time in ℓ) to see that it approaches

$$\sum_{k=1}^{\infty} \prod_{j=1}^{\infty} \mu_j(E_j^k) = \sum_{k=1}^{\infty} \mu_0 \left(\prod_{j=1}^{\infty} E_j^k \right)$$

as desired. Hence μ_0 extends to a premeasure on $\mathcal{A}$, and therefore to a measure on the sigma-algebra generated by $\mathcal{A}$ by Theorem 1.5. $\square$

Exercise 16: Consider the d-dimensional torus $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$. Identify $\mathbb{T}^d$ as $\mathbb{T}^1 \times \cdots \times \mathbb{T}^1$ and let μ be the product measure on $\mathbb{T}^d$ given by $\mu = \mu_1 \times \cdots \times \mu_d$, where μ_j is Lebesgue measure on X_j identified with the circle $\mathbb{T}$. That is, if we represent each point in X_j uniquely as k_j with $0 < x_j \leq 1$, then the measure μ_j is the induced Lebesgue measure on $\mathbb{R}$ restricted to $(0, 1]$.

(a) Check that the completion μ is Lebesgue measure induced on the cube $Q = \{x : 0 < x_j \leq 1, j = 1, \dots, d\}$.

(b) For each function f on Q let $\tilde{f}$ be its extension to $\mathbb{R}^d$ which is periodic, that is, $\tilde{f}(x + z) = \tilde{f}(x)$ for every $z \in \mathbb{Z}^d$. Then f is measurable on $\mathbb{T}^d$ if and only if $\tilde{f}$ is measurable on $\mathbb{R}^d$, and f is continuous on $\mathbb{T}^d$ if and only if $\tilde{f}$ is continuous on $\mathbb{R}^d$.

(c) Suppose f and g are integrable on $\mathbb{T}^d$. Show that the integral defining $(f*g)(x) = \int_{\mathbb{T}^d} f(x-y)g(y)dy$ is finite for a.e. x, that $f*g$ is integrable over $\mathbb{T}^d$, and that $f * g = g * f$.

(d) For any integrable function f on $\mathbb{T}^d$, write

$$f \sim \sum_{n \in \mathbb{Z}^d} a_n e^{2\pi i n \dot{x}}$$

to mean that $a_n = \int_{\mathbb{T}^d} f(x) e^{-2\pi i n \dot{x}} dx$. Prove that if g is also integrable, and $g \sim \sum_{n\in\mathbb{Z}^d} b_n e^{2\pi i n\dot{x}}$, then

$$f * g \sim \sum_{n\in\mathbb{Z}^d} a_n b_n e^{2\pi i n \dot{x}}.$$

(e) Verify that $\{e^{2\pi i n\dot{x}}\}$ is an orthonormal basis for $L^2(\mathbb{T}^d)$. As a result $\|f\|_{L^2(\mathbb{T}^d)} = \sum_{n\in\mathbb{Z}^d} |a_n|^2$.

(f) Let f be any continuous periodic function on $\mathbb{T}^d$. Then f can be uniformly approximated by finite linear combinations of the exponentials $\{e^{2\pi i n\dot{x}}\}_{n\in\mathbb{Z}^d}$.

Solution.

(a) This follows from the translation invariance of μ. To show that the product of translation invariant measures is translation invariant, let $E = E_1 \times \cdots \times E_d$ be a measurable rectangle in $\mathbb{T}^d$, and $x = (x_1, \ldots, x_d) \in \mathbb{T}^d$. Then

$$\mu(E+x) = \mu\big((E_1+x_1)\times\cdots\times(E_d+x_d)\big) = \mu_1(E_1+x_1)\ldots\mu_d(E_d+x_d) = \mu_1(E_1)\ldots\mu_d(E_d) = \mu(E)$$

so μ is translation invariant on measurable rectangles. This implies that the outer measure μ_* generated by coverings of measurable rectangles is also translation invariant. But μ is just the restriction of μ_* to the sigma-algebra of Carathéodory-measurable sets, so it is translation invariant as well. This implies that μ is a multiple of Lebesgue measure; since $\mu(\mathbb{T}^d) = m(Q) = 1$, we must have $m = \mu$ (modulo the correspondence between Q and $\mathbb{T}^d$).

(b) This is blindingly obvious, but

$$\begin{aligned} f \text{ m'ble (resp. cts)} &\Leftrightarrow f^{-1}(U) \text{ m'ble (resp. open) for open } U \\ &\Leftrightarrow \tilde{f}^{-1}(U) \text{ m'ble (resp. open) in } \mathbb{R}^d \\ &\Leftrightarrow \tilde{f} \text{ m'ble (resp. cts).} \end{aligned}$$

Here we use the fact that $\tilde{f}^{-1}(U)$ is a lattice consisting of translates of $f^{-1}(U)$ by $\mathbb{Z}^d$; such a set is open or measurable iff $f^{-1}(U)$ is.

(c) This is a simple application of Tonelli's Theorem, exactly analogous to the case in $L^1(\mathbb{R}^d)$:

$$\begin{aligned} \int_{\mathbb{T}^d} |f*g(x)|dx &= \int_{\mathbb{T}^d} \left| \int_{\mathbb{T}^d} f(x-y)g(y)dy \right| dx \\ &\le \int_{\mathbb{T}^d}\int_{\mathbb{T}^d} |f(x-y)||g(y)|dydx \\ &= \int_{\mathbb{T}^d}\int_{\mathbb{T}^d} |f(x-y)||g(y)|dxdy \\ &= \|f\|_{L^1(\mathbb{T}^d)}\|g\|_{L^1(\mathbb{T}^d)} \end{aligned}$$

so $f*g$ is integrable on $\mathbb{T}^d$. This in turn implies that it is finite a.e. Finally, the change of variables $u = x - y$ shows that

$$f*g(x) = \int f(x-y)g(y)dy = \int f(u)g(x-u)du = g*f(x).$$

(d) Once again, there is absolutely nothing different from the one-variable case. Since $f * g$ is integrable by our above remarks, and $|e^{-2\pi i n\cdot x}| = 1$, $f * g(x)e^{-2\pi i n\cdot x}$ is also integrable, so by Fubini's theorem

$$\begin{aligned}\int_{\mathbb{T}^d} f * g(x)e^{-2\pi i n\cdot x}dx &= \int_{\mathbb{T}^d} e^{-2\pi i n\cdot x}\int_{\mathbb{T}^d} f(x-y)g(y)dydx \\ &= \int_{\mathbb{T}^d} g(y)\int_{\mathbb{T}^d} f(x-y)e^{-2\pi i n\cdot x}dxdy \\ &= \int_{\mathbb{T}^d} g(y)e^{-2\pi i n\cdot y}\int_{\mathbb{T}^d} f(x-y)e^{-2\pi i n\cdot(x-y)}dxdy \\ &= \int_{\mathbb{T}^d} g(y)e^{-2\pi i n\cdot y}a_n dy \\ &= a_n b_n.\end{aligned}$$

(e) The orthonormality of this system is evident, since

$$\int_{\mathbb{T}^d} e^{-2\pi i n\cdot x}e^{2\pi i m\cdot x}dx = \int_{\mathbb{T}^d} e^{2\pi i(m-n)\cdot x}dx = \prod_{j=1}^{d}\int_{\mathbb{T}} e^{2\pi i(m_j-n_j)x_j}dx_j = \prod_{j=1}^{d}\delta_{m_j}^{n_j} = \delta_m^n$$

where δ is the Kronecker delta function. To show completeness, we use the fact that an orthonormal system $\{e_n\}$ in a Hilbert space is complete iff $\langle f, e_n\rangle = 0$ for all $n \Rightarrow f = 0$. Suppose

$$0 = \langle f, e^{2\pi i n\cdot x}\rangle = \int_{\mathbb{T}^d} f(x)e^{-2\pi i n\cdot x}dx = \int_{\mathbb{T}} e^{-2\pi i n_1 x_1}\int_{\mathbb{T}} e^{-2\pi i n_2 x_n}\cdots\int_{\mathbb{T}^d} e^{-2\pi i n_d x_d}f(x_1,\ldots,x_d)dx_d\ldots dx_1$$

for all $n_1,\ldots,n_d$, where the use of Fubini's theorem is justified by the integrability of f and the fact that $|e^{-2\pi i n_k x_k}| = 1$. Let

$$F_1(x_1) = \int_{\mathbb{T}} e^{-2\pi i n_2 x_2}\int_{\mathbb{T}} e^{-2\pi i n_3 x_3}\cdots\int_{\mathbb{T}} e^{-2\pi i n_d x_d}f(x_1,\ldots,x_d)dx_d\ldots dx_2.$$

Then $\int_{\mathbb{T}} e^{-2\pi i n_1 x_1}F_1(x_1)dx_1 = 0$ for all n_1, so $F_1(x_1) = 0$ a.e. by the completeness of exponentials in the 1-dimensional case. Now let

$$F_2(x_1,x_2) = \int_{\mathbb{T}} e^{-2\pi i n_3 x_3}\cdots\int_{\mathbb{T}} e^{-2\pi i n_d x_d}f(x_1,\ldots,x_d)dx_d\ldots dx_1.$$

For any fixed value of x_1, $F_n(x_1,x_2)$ is a function of x_2 with the property that $\int_{\mathbb{T}} e^{-2\pi i n_2 x_n}F_2(x_1,x_2)dx_2 = F_1(x_1) = 0$ a.e., so we must have $F_2(x_1,x_2) = 0$ for a.e. x_2. Continuing inductively, we see that $f(x_1,\ldots,x_d) = 0$ for a.e. $x_1,\ldots,x_d$.

(f) This problem is begging for the Stone-Weierstrass theorem, but since we haven't covered that in class, I'll reluctantly do the convolution stuff. Let $g_\delta(x) = \frac{1}{\delta^d}\chi_{[0,\delta]^d}$. Then $\int g_\delta(x)dx = 1$, so

$$\begin{aligned}|f(x) - f * g_\delta(x)| &= \left|f(x)\int g_\delta(y)dy - \int f(x-y)g_\delta(y)\right|dy \\ &\le \int |f(x) - f(x-y)||g_\delta(y)|dy.\end{aligned}$$

By the uniform continuity of f, given $\epsilon > 0$ there exists $\delta_0 > 0$ such that $|x-y| < \delta_0 \Rightarrow |f(x) - f(y)| < \epsilon$. For $\delta < \delta_0$,

$$|f(x) - f * g_\delta(x)| \le \int \epsilon|g_\delta(y)|dy = \epsilon.$$

Thus, $f * g_\delta(x) \to f(x)$ uniformly as $\delta \to 0$. However, if a_n are the Fourier coefficients of f and b_n^δ those of g_δ, then $\sum |a_n|^2 < \infty$ and $\sum |b_n^\delta|^2 < \infty$ because f and g_δ are in L^2 so their Fourier transforms are as well. Then by the Cauchy-Schwarz inequality, $\sum |a_n b_n^\delta| < \infty$. This implies that $\widehat{f * g_\delta} \in L^1$ and since $f * g_\delta \in L^1$, Fourier inversion holds and we have $f * g_\delta(x) = \sum a_n b_n e^{2\pi i n \cdot x}$ a.e. (in fact, everywhere, since both sides are continuous). Now we can choose δ such that $|f - f * g_\delta| < \frac{\epsilon}{2}$ everywhere. For this δ, since the tails of the convergent series $\sum a_n b_n^\delta$ go to zero, we can choose some truncation $\sum_{|n| \le N} |a_n b_n|$ such that $\sum_{|n|>N} |a_n b_n^\delta| < \frac{\epsilon}{2}$. Then for any x,

$$\begin{aligned}
\left| f(x) - \sum_{|n| \le N} a_n b_n e^{2\pi i n \cdot x} \right| &\le |f(x) - f * g_\delta(x)| + \left| f * g_\delta(x) - \sum_{|n| \le N} e^{2\pi i n \cdot x} \right| \\
&= |f(x) - f * g_\delta(x)| + \left| \sum_{|n>N|} a_n b_n^\delta e^{2\pi i n \cdot x} \right| \\
&\le |f(x) - f * g_\delta(x)| + \sum_{|n>N|} |a_n b_n^\delta| \\
&\le \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.
\end{aligned}$$

Thus, f can be uniformly approximated by trigonometric polynomials. □

Exercise 17: By reducing to the case $d = 1$, show that each "rotation" $x \mapsto x + \alpha$ of the torus $\mathbb{T}^d = \mathbb{R}^d / \mathbb{Z}^d$ is measure preserving, for any $\alpha \in \mathbb{R}^d$.

Solution. We first suppose that $E \subset \mathbb{T}^d$ is a measurable rectangle $E = E_1 \times \cdots \times E_d$ where $E_k \subset \mathbb{T}$ for $k = 1, \ldots, d$. Then

$$m(\tau^{-1}(E)) = m(E - \alpha) = m((E_1 - \alpha_1) \times \cdots \times (E_d - \alpha_d)) = m(E_1) \ldots m(E_d) = m(E).$$

Hence τ is measure-preserving on measurable rectangles. But since the measure of any set is computed in terms of its coverings by measurable rectangles, τ is measure-preserving on all measurable sets. (We actually use here the fact that τ^{-1} is measure-preserving as well; if $\{R_k\}$ is a covering of E by measurable rectangles, then $\{\tau^{-1}(R_k)\}$ is a covering of $\tau^{-1}(E)$ with the same measure; conversely, if $\{R'_k\}$ is a covering of $\tau^{-1}(E)$ by measurable rectangles, then $\{\tau(R'_k)\}$ is a covering of E with the same measure.) □

Exercise 18: Suppose τ is a measure-preserving transformation on a measure space (X, μ) with $\mu(X) = 1$. Recall that a measurable set E is invariant if $\tau^{-1}(E)$ and E differ by a set of measure zero. A sharper notion is to require that $\tau^{-1}(E)$ equal E. Prove that if E is any invariant set, there is a set E' so that $E' = \tau^{-1}(E)$, and E and E' differ by a set of measure zero.

Solution. Let

$$E' = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \tau^{-k}(E).$$

Then

$$E' \setminus E = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \tau^{-k}(E) \setminus E$$

and

$$E \setminus E' = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} E \setminus \tau^{-k}(E).$$

But $E \setminus \tau^{-k}(E)$ and $\tau^{-k}(E) \setminus E$ both have measure zero (this follows from $m(E\Delta E') = 0$ by an easy induction), and countable unions and intersections of null sets are null, so $m(E\Delta E') = 0$. Moreover,

$$\tau^{-1}(E') = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \tau^{-1}\tau^{-k}(E) = \bigcap_{n=2}^{\infty} \bigcup_{k=n}^{\infty} \tau^{-k}(E) = E'$$

because the sets inside the intersection are nested so we get the same set whether we start at $n = 1$ or $n = 2$. □

Exercise 19: Let τ be a measure-preserving transformation on (X, μ) with $\mu(X) = 1$. Then τ is ergodic if and only if whenever ν is absolutely continuous with respect to μ and ν is invariant (that is, $\nu(\tau^{-1}(E)) = \nu(E)$ for all measurable sets E), then $\nu = c\mu$, with c a constant.

Solution. We use the fact that τ is ergodic iff the only functions with $f \circ \tau = f$ a.e. are constant a.e., as well as the fact that $\int_E f d\mu = \int_{\tau^{-1}(E)} f \circ \tau d\mu$ for measure-preserving maps τ. Let τ be ergodic and let $\nu \ll \mu$ be an invariant measure. By the Radon-Nikodym theorem, $d\nu = h d\mu$ for some function $h \in L^1(\mu)$. Then for any measurable E,

$$\nu(E) = \int_E h d\mu = \int_{\tau^{-1}(E)} h \circ \tau d\mu.$$

By the invariance of τ, this equals

$$\nu(\tau^{-1}(E)) = \int_{\tau^{-1}(E)} h d\mu.$$

Since this is true for any measurable E, $h \circ \tau = h$ a.e. But since τ is ergodic, this implies h is constant a.e., so $\nu = c\mu$ for some constant c. Conversely, suppose every invariant absolutely continuous measure ν is a constant times μ. Let $f \in L^1$ be any function with the property that $f \circ \tau = f$ a.e. Define an absolutely continuous measure ν by $d\nu = f d\mu$. Then the above calculation (run in reverse) shows that ν is invariant, so $\nu = c\mu$. But this implies that $f d\mu = d\nu = c d\mu$ so f is constant a.e. Hence τ is ergodic. □

Exercise 20: Suppose τ is a measure-preserving transformation on (X, μ). If

$$\mu(\tau^{-n}(E) \cap F) \to \mu(E)\mu(F)$$

as $n \to \infty$ for all measurable sets E and F, then $(T^n f, g) \to (f, 1)(1, g)$ whenever $f, g \in L^2(X)$ with $(Tf)(x) = f(\tau(x))$. Thus τ is mixing.

Solution. Suppose $\mu(\tau^{-n}(E)\cap F)\to\mu(E)\mu(F)$ for measurable E and F. This means that $(T^n f,g)\to(f,1)(1,g)$ if f and g are characteristic functions: say $f=\chi_E$ and $g=\chi_F$, then

$$(T^n f,g)=\int\chi_E(\tau^n(x))\chi_F(x)dx=\int\chi_{\tau^{-n}(E)}\chi_F=m(\tau^{-n}(E)\cap F)\to\mu(E)\mu(F)=(f,g).$$

By linearity in f and conjugate-linearity in g, this implies $(T^n f,g)\to(f,1)(1,g)$ if f and g are measurable simple functions. Now let $f,g\in L^2$ and let $f_m\to f$ and $g_m\to g$ be sequences of measurable simple functions. For each m we have

$$(T^n f,g)=(T^n f,g-g_m)+(T^n(f-f_m),g)+(T^n f_m,g_m).$$

Since T is an isometry, $|(T^n f,g-g_m)|\le\|T^n f\|\|g-g_m\|=\|f\|\|g-g_m\|\to 0$ as $m\to\infty$. Similarly, $(T^n(f-f_m),g)\to 0$ uniformly in n as $m\to\infty$. Finally, $(T^n f_m,g_m)\stackrel{n}{\to}(f_m,1)(1,g_m)\stackrel{m}{\to}(f,1)(1,g)$. (To be more precise about this business of taking limits in two different variables, choose f_m and g_m with $\|f_m-f\|,\|g_m-g\|<\epsilon$. Then since T is an isometry,

$$(T^n f,g)=(T^n f_m,g_m)+h(n)$$

where $\|h(n)\|<2\epsilon$ for all n. Letting $n\to\infty$, we see that $(T^n f,g)$ is eventually within 2ϵ of $(f_m,1)(1,g_m)$, which in turn is within $C\epsilon$ of $(f,1)(1,g)$ for some constant C. This is true for all ϵ, so $(T^n f,g)\to(f,1)(g,1)$.) $\square$

Exercise 21: Let $\mathbb{T}^d$ be the torus, and $\tau:x\mapsto x+\alpha$ the mapping arising in Exercise 17. Then τ is ergodic if and only if $\alpha=(\alpha_1,\dots,\alpha_d)$ with $\alpha_1,\dots,\alpha_d$, and 1 are linearly independent over the rationals. To do this show that:

(a)
$$\frac{1}{m}\sum_{k=0}^{m-1}f(\tau^k(x))\to\int_{\mathbb{T}^d}f(x)dx$$
as $n\to\infty$, for each $x\in\mathbb{T}^d$, whenever f is continuous and periodic and α satisfies the hypothesis.

(b) Prove as a result that in this case τ is uniquely ergodic.

Solution.

(a) Suppose first that $\alpha_1,\dots,\alpha_d$ and 1 are dependent over $\mathbb{Q}$, say
$$a_1\alpha_1+\cdots+a_d\alpha_d=\frac{p}{q},$$
where $\alpha_i\in\mathbb{Q}$. Let
$$E=\{(x_1,\dots,x_d):0<\{q(a_1x_1+\cdots+a_dx_d)\}<\frac{1}{2}\},$$
where $\{z\}=z-\lfloor z\rfloor$ denotes the fractional part of z. Then $m(E)=\frac{1}{2}$ but $E=\tau^{-1}(E)$, so τ is not ergodic.

On the other hand, suppose $\alpha_1,\dots,\alpha_d$ and 1 are independent over $\mathbb{Q}$. Let $f(x)=e^{2\pi in\cdot x}$ be any complex exponential. If $n=0$, then
$$\frac{1}{m}\sum_{k=0}^{m-1}f((\tau^k(x))=1=\int_{\mathbb{T}^d}f(x)dx.$$

If $n \neq 0$, then

$$\frac{1}{m}\sum_{k=0}^{m-1} f(x)dx = \frac{1}{m}e^{2\pi i n\cdot x}\sum_{k=0}^{m-1} e^{2\pi i k n\cdot\alpha} = \frac{e^{2\pi i n\cdot x}}{m}\frac{1-e^{2\pi i m n\cdot\alpha}}{1-e^{2\pi i n\cdot\alpha}}.$$

Since $|1-e^{2\pi i m n\cdot\alpha}| \leq 2$, this goes to zero as $m \to \infty$, so

$$\frac{1}{m}\sum_{k=0}^{m-1} f((\tau^k(x)) \to \int_{\mathbb{T}^d} f(x)dx.$$

Finally, since complex exponentials are uniformly dense in the continuous periodic functions by exercise 16f, the above limit holds for any continuous periodic function: Let f be a continuous periodic function and P_ϵ a finite linear combination of complex exponentials with $|f - P_\epsilon| < \epsilon$ everywhere. Choose n sufficiently large that $|A_m P_\epsilon - \int P_\epsilon dx| < \epsilon$ for all $m > n$, where $A_m g(x) = \frac{1}{m}\sum_{k=0}^{m-1} g(\tau^k(x))$. Then for $m > n$,

$$\left|A_m f(x) - \int f(x)dx\right| \leq |A_m f(x) - A_m P_\epsilon(x)| + \left|A_m P_\epsilon(x) - \int P_\epsilon(x)dx\right| + \left|\int (P_\epsilon(x) - f(x))dx\right|$$
$$< \epsilon + \epsilon + \epsilon = 3\epsilon.$$

(b) Unique ergodicity follows by the same logic as the 1-variable case. Let ν be any invariant measure; then part (a) plus the Mean Ergodic Theorem shows that $P_\nu(f) = \int f dx$, where P_ν is the projection in $L^2(\nu)$ onto the subspace of invariant functions. This implies that the image of P_ν is just the constant functions. But we know that the $L^2(\nu)$ projection of f onto the constants is $\int f d\nu$, so we must have $\int f dx = \int f d\nu$ for continuous f. Since characteristic functions of open rectangles can be L^2-approximated by continuous functions, this implies that $m(R) = \nu(R)$ for any open rectangle R. But this implies that m and ν agree on the Borel sets. Hence m is uniquely ergodic for this τ.

□

Exercise 26: There is an L^2 version of the maximal ergodic theorem. Suppose τ is a measure-preserving transformation on (X, μ). Here we do not assume that $\mu(X) < \infty$. Then

$$f^*(x) = \sup \frac{1}{m}\sum_{k=0}^{m-1} |f(\tau^k(x))|$$

satisfies

$$\|f^*\|_{L^2(X)} \leq c\|f\|_{L^2(X)}, \quad \text{whenever } f \in L^2(X).$$

The proof is the same as outlined in Problem 6, Chapter 5 for the maximal function on $\mathbb{R}^d$. With this, extend the pointwise ergodic theorem to the case where $\mu(X) = \infty$, as follows:

(a) Show that $\lim_{m\to\infty} \frac{1}{m}\sum_{k=0}^{m-1} f(\tau^k(x))$ converges for a.e. x to $P(f)(x)$ for every $f \in L^2(X)$, because this holds for a dense subspace of $L^2(X)$.

(b) Prove that the conclusion holds for every $f \in L^1(X)$, because it holds for the dense subspace $L^1(X) \cap L^2(X)$.

Solution.

(a) We use the subspaces $S = \{f \in L^2 : f \circ \tau = f\}$ and $S_1 = \{g - Tg : g \in L^2\}$ from the proof of the mean ergodic theorem. As shown there, $L^2(X) = S \oplus \bar{S}_1$. Given $f \in L^2$, let $\epsilon > 0$ and write $f = f_0 + f_1 + f_2$ where $f_0 \in S$, $f_1 + f_2 \in \bar{S}_1$, $f_1 \in S_1$, $\|f_2\| < \epsilon$. Since $f_1 \in S_1$, $f_1 = g - Tg$ for some $g \in L^2$. Let $h = f_0 + f_1$. Then $A_m f_0 = f_0 = P f_0$ for all m, and

$$A_m f_1 = \frac{1}{m} \sum_{k=0}^{m-1} T^k (g - Tg) = \frac{1}{m}(g - T^m g).$$

Clearly $\frac{1}{m} g(x) \to 0$ for all x as $m \to \infty$. Moreover, as shown on page 301, $\frac{1}{m} T^m g(x) \to 0$ for almost all x; one can see this from the fact that, by the monotone convergence theorem,

$$\int_X \sum_{m=1}^{\infty} \frac{1}{m^2} |T^m(g)(x)|^2 = \sum_{m=1}^{\infty} \frac{1}{m^2} \int_X |T^m(g)(x)|^2 = \sum_{m=1}^{\infty} \frac{1}{m^2} \|T^m g\|^2 = \|g\| \sum_{m=1}^{\infty} \frac{1}{m^2} < \infty.$$

Since $\sum \frac{1}{m^2} |T^m(g)(x)|^2$ is integrable, it is finite almost everywhere, which means the terms in the series tend to zero for almost all x. The upshot is that $A_m f_1(x) \to P f_1(x)$ for a.a. x, so that $A_m h(x) \to Ph(x)$ a.e. Finally, let

$$E_\alpha = \left\{ x \in X : \lim_{m \to \infty} \sup |A_m f(x) - Pf(x)| > \alpha \right\}.$$

Since $A_m f - Pf = A_m h - Ph + A_m(f - h) - P(f - h)$, $E_\alpha \subset N \cup F_\alpha \cup G_\alpha$, where $N = \{x : A_m h(x) \not\to Ph(x)\}$,

$$F_\alpha = \left\{ x \in X : (f - h)^*(x) > \frac{\alpha}{2} \right\},$$

and

$$G_\alpha = \left\{ x \in X : |P(f - g)(x)| > \frac{\alpha}{2} \right\}.$$

Now $\mu(N) = 0$ as we have already established, and by the L^2 maximal theorem, $\mu(F_\alpha) < \frac{\|(f-h)^*\|_2^2}{(\alpha/2)^2} < \frac{4c^2\epsilon^2}{\alpha^2}$; similarly, since $\|Py\| \le \|y\|$ for $y \in L^2$, $\mu(G_\alpha) < \frac{\|f-h\|_2^2}{(\alpha/2)^2} < \frac{4\epsilon^2}{\alpha^2}$. Since ϵ is arbitrary, $\mu(F_\alpha) = \mu(G_\alpha) = 0$ for all $\alpha > 0$. Thus, $\mu(E_\alpha) = 0$ for all $\alpha > 0$, so $A_m f(x) \to Pf(x)$ a.e.

(b) Let $f \in L^1$. For any $\epsilon > 0$, choose $g \in L^2 \cap L^1$ with $\|f - g\|_1 < \epsilon$. Then

$$A_m f(x) - Pf(x) = A_m g(x) - Pg(x) + A_m(f - g)(x) - P(f - g)(x).$$

Let

$$E_\alpha = \{x \in X : \limsup |A_m f(x) - Pf(x)| > 2\alpha\}.$$

Then $E_\alpha \subset N \cup F_\alpha \cup G_\alpha$ where $N = \{x : A_m g(x) - Pg(x) \not\to 0\}$,

$$F_\alpha = \{x \in X : \limsup |A_m(f - g)(x)| > \alpha\},$$

and

$$G_\alpha = \{x \in X : \limsup |P(f - g)(x)| > \alpha\}.$$

Now $\mu(N) = 0$ by part (a), and by inequality (24) on page 297,

$$\mu(F_\alpha) \leq \frac{A}{\alpha}\|f-g\|_1 < \frac{A}{\alpha}\epsilon.$$

Also $\|P(f-g)\| \leq \|f-g\| < \epsilon$, so by Chebyshev's inequality,

$$\mu(G_\alpha) \leq \frac{1}{\alpha}\|f-g\|_1 < \frac{1}{\alpha}\epsilon.$$

Since ϵ is arbitrary, this implies $\mu(F_\alpha) = \mu(G_\alpha) = \mu(E_\alpha) = 0$ for all $\alpha > 0$. Hence $A_m f(x) \to Pf(x)$ for a.a. x.

□

Exercise 27: We saw that if $\|f_n\|_{L^2} \leq 1$, then $\frac{f_n(x)}{n} \to 0$ as $n \to \infty$ for a.e. x. However, show that the analogue where one replaces the L^2-norm by the L^1-norm fails, by constructing a sequence $\{f_n\}$, $f_n \in L^1(X)$, $\|f_n\|_{L^1} \leq 1$, but with $\limsup \frac{f_n(x)}{n} = \infty$ for a.e. x.

Solution. This is yet another example of why Stein & Shakarchi sucks. The problem doesn't say anything about conditions on X. In fact, the hint seems to assume that $X = [0,1]$. I will assume that X is σ-finite. I will also assume that the measure μ has the property that for any measurable set E and any real number α with $0 \leq \alpha \leq \mu(E)$, there is a subset $S \subset E$ with $\mu(S) = \alpha$. (This property holds, for example, in the case of Lebesgue measure, or any measure which is absolutely continuous with respect to Lebesgue measure. In fact, we don't need quite this stringent a requirement–it isn't necessary that *every* subset of X have this nice property, but only that we can find a nested sequence of subsets of X_n whose measures we can control this way, where $X = \cup X_n$ and $\mu(X_n) < \infty$.)

Given these assumptions, let $X = \cup X_n$ where $\mu(X_n) < \infty$. We construct a sequence E_n of measurable subsets with the properties that $\mu(E_n) \leq \frac{1}{n \log n}$ and that every $x \in X$ is in infinitely many E_n. To do this, we will construct countably many finite sequences and then string them all together. The first sequence $E_2, \ldots, E_{N_1}$ will have the property that $X_1 = \cup_{j=2}^{N} E_j$, and $\mu(E_j) \leq \frac{1}{j \log j}$. To do this, let E_2 be any subset of X_1 with measure $\frac{1}{2\log 2}$, unless $\mu(X_1) \leq \frac{1}{2\log 2}$, in which case $E_2 = X_1$. Let E_3 be any subset of $X_1 \setminus E_2$ with measure $\frac{1}{3\log 3}$, unless $\mu(X_1 \setminus E_2) \leq \frac{1}{3\log 3}$, in which case $E_3 = X_1 \setminus E_2$. Let E_4 be a subset of $X_1 \setminus (E_2 \cup E_3)$ with measure $\frac{1}{4\log 4}$, or $X_1 \setminus (E_2 \cup E_3)$ if this has measure at most $\frac{1}{4\log 4}$. This process will terminate in finitely many steps because $\sum \frac{1}{n\log n}$ diverges and $\mu(X_1)$ is finite.

We then construct a second finite sequence of sets $E_{N_1+1}, \ldots, E_{N_2}$ whose union is $X_1 \cup X_2$, a third finite sequence whose union is $X_1 \cup X_2 \cup X_3$, etc. Let E_n be the concatenation of all these finite sequences. Then every point in X is in infinitely many E_n, and $\mu(E_n) \leq \frac{1}{n\log n}$.

Now let $f_n = n \log n \chi_{E_n}$. Then $\|f_n\|_1 = n \log n \mu(E_n) \leq 1$. However, $\frac{f_n(x)}{n} = \log n \chi_{E_n}(x)$ and since x is in infinitely many E_n, $\limsup \frac{f_n(x)}{n} = \infty$ for all x. □

Exercise 28: We know by the Borel-Cantelli lemma that if $\{E_n\}$ is a collection of measurable sets in a measure space (X, μ) and $\sum_{n=1}^{\infty} \mu(E_n) < \infty$, then $E = \limsup\{E_n\}$ has measure zero.

In the opposite direction, if τ is a mixing measure-preserving transformation on X with $\mu(X) = 1$, then whenever $\sum_{n=1}^{\infty} \mu(E_n) = \infty$, there are integers $m = m_n$ so that if $E'_n = \tau^{-m_n}(E_n)$, then $\limsup(E'_n) = X$ except for a set of measure 0.

Solution. Let $F_n = E_n^c$. Since $\sum \mu(E_n) = \infty$, a theorem on infinite products (e.g. Corollary 5.6 on page 166 of Conway, *Functions of One Complex Variable*) says that $\prod \mu(F_n) = 0$. Let $F'_n = F_1$. Because τ is mixing, there exists m_2 such that

$$\mu(\tau^{-m_2}(F_2) \cap F_1) < \mu(F_1)\mu(F_2) + \frac{1}{2^2}.$$

and let $F'_2 = \tau^{-m^n}(F_2)$. Next, choose m_3 such that

$$\mu(\tau^{-m_3}(F_3) \cap (F'_1 \cap F'_2)) < \mu(F'_1 \cap F'_2)\mu(F_3) + \frac{1}{2^3}$$

and

$$\mu(\tau^{-m_3}(F_3) \cap F'_2) < \mu(F'_2)\mu(F_3) + \frac{1}{2^3}$$

and let $F'_3 = \tau^{-m_3}(F_3)$. Note that this implies

$$\mu(F'_1 \cap F'_2 \cap F'_3) < \mu(F_3)\left(\mu(F_1)\mu(F_2) + \frac{1}{2^2}\right) + \frac{1}{2^4} < \mu(F_1)\mu(F_2)\mu(F_3) + \frac{1}{2^2}\mu(F_3) + \frac{1}{2^3}$$

and

$$\mu(F'_2 \cap F'_3) < \mu(F_2)\mu(F_3) + \frac{1}{2^3}.$$

Continuing, we choose m_k such that

$$\mu\left(\tau^{-m_k}(F_k) \cap \bigcap_{j=\ell}^{k-1} F'_j\right) < \mu(F_k)\mu\left(\bigcap_{j=\ell}^{k-1} F'_j\right) + \frac{1}{2^k}$$

for all $\ell = 1, \ldots, k-1$; by induction this is less than

$$\prod_{j=\ell}^{k} \mu(F_j) + \sum_{j=\ell+1}^{k} \frac{1}{2^j} \prod_{i=j+1}^{k} \mu(F_i)$$

where the latter product is taken to be 1 if empty (i.e. when $j = k$). I will prove that this tends to 0 as $k \to \infty$, for any fixed ℓ. As noted before, $\prod_{j=\ell}^{\infty} \mu(F_j) = 0$ (any tail of the sum diverges, so any tail of the product tends to zero) so the first term tends to zero. Similarly, $\prod_{j=i}^{\infty} \mu(F_j) = 0$ for any fixed i. Thus, if we split the sum into a sum from $\ell + 1$ to M and a sum from M to k, the finitely many terms from $\ell + 1$ to M can all be made arbitrarily small because they each approach zero as $k \to \infty$. On the other hand, each term $\frac{1}{2^j} \prod_{i=j+1}^{k} \mu(F_i)$ is at most $\frac{1}{2^j}$, so their sum is less than $\sum_{j=M}^{\infty} \frac{1}{2^j}$ which can be made arbitrarily small by an appropriate choice of M. This proves that

$$\mu\left(\bigcap_{j=\ell}^{\infty} F'_j\right) = 0$$

for all ℓ. Then

$$\mu\left(\bigcup_{\ell=1}^{\infty}\bigcap_{j=\ell}^{\infty}F'_j\right)=0.$$

But the complement of this set is just $\limsup E'_j$ where $E'_j=\tau^{-m_j}(E_j)$. Thus, $\limsup E'_j$ is almost all of X. $\square$

Chapter 6.8, Page 319

Problem 1: Suppose Φ is a C^1 bijection of an open set $\mathcal{O}$ in $\mathbb{R}^d$ with another open set $\mathcal{O}'$ in $\mathbb{R}^d$.

(a) If E is a measurable subset of $\mathcal{O}$, then $\Phi(E)$ is also measurable.
(b) $m(\Phi(E))=\int_E|\det\Phi'(x)|dx$, where Φ' is the Jacobian of Φ.
(c) $\int_{\mathcal{O}'}f(y)dy=\int_{\mathcal{O}}f(\Phi(x))|\det\Phi'(x)|dx$ whenever f is integrable on $\mathcal{O}'$.

Solution.

(a) If K is compact, then $\Phi(K)$ is also compact by continuity. Now if A is F_σ, then A is σ-compact (since every closed set in $\mathbb{R}^n$ is a countable union of compact sets), so $\Phi(A)$ is also σ-compact and hence F_σ. Thus, Φ maps F_σ sets to F_σ sets. Since measurable sets are precisely those that differ from F_σ sets by a set of measure zero, it suffices to show that Φ maps sets of measure zero to sets of measure zero. (The following argument is adopted from Rudin p. 153.) Let $E\subset\mathcal{O}$ have measure zero. For each integer n, define

$$F_n=\{x\in\mathcal{E}:|\Phi'(x)|<n\}.$$

Then for any $x\in F_n$, the definition of derivative implies that $\frac{|\Phi(y)-\Phi(x)|}{|y-x|}<n$ for $|x-y|<\delta$ for some $\delta>0$. Define $F_{n,p}\subset F_n$ to be the set of x for which $\delta=\frac{1}{p}$ works. Now $m(F_{n,p})=0$ because it's a subset of E. I claim that we can cover $F_{n,p}$ by balls of radius less than $\frac{1}{p}$ with centers in $F_{n,p}$ and total measure at most ϵ. To do this, we can first cover $F_{n,p}$ by an open set of arbitrarily small measure. This open set can be decomposed into cubes of arbitrarily small diameter, as shown in chapter 1. If the diameter is sufficiently small (less than a constant times $\frac{1}{p}$), we can cover whichever of these cubes intersect $F_{n,p}$ with a ball centered at a point of $F_{n,p}$ and radius less than $\frac{1}{p}$; the other cubes we discard. The total measure of the resulting balls is at most a constant times the measure of the open set. (This constant comes from finding the maximal possible ratio of volumes of the ball covering one of these small-diameter cubes to the cube, which comes when the center of the ball is at one corner of the cube and the radius of the ball is $\sqrt{2}$ times the cube's diameter.) This proves the claim. Now if we cover $F_{n,p}$ by such cubes B_j, centered at x_j and with radius $r_j<\frac{1}{p}$ then for $x\in B_j$ we have $|T(x)-T(x_j)|\le n|x-x_j|\le nr_j$. This implies that

$$T(F_{n,p})\subset\bigcup_j B_{nr_j}(x_j)\Rightarrow m(T(F_{n,p})\le\sum m(B_{nr_j}(x_j))=n^d\sum m(B_{r_j}(x_j))<n^d\epsilon.$$

Since ϵ was arbitrary, this shows that $m(T(F_{n,p})) = 0$. Since $\mathcal{E} = \cup_{n,p} F_{n,p}$, this implies $m(\Phi(\mathcal{E})) = 0$.

(b) I will prove this in the case where E is a closed rectangle contained in $\mathcal{O}$, and then show that this implies the general case.
Thus, assume $E \subset R \subset U$ for a compact rectangle R. By absolute continuity of each of the n components of Φ', given any $\epsilon > 0$ there exists $\delta > 0$ such that $|x - y| < \delta$, for $x, y \in R$, implies $|(\Phi'(x) - \Phi'(y))_j| < \epsilon$ for $j = 1, \ldots, n$. Divide R into cubes Q_k of diameter less than δ. Let a_k be the center of Q_k. Then for $x \in Q_k$, the mean value theorem implies $\Phi(x)_j - \Phi(a_k)_j = \Phi'(c)(x - a_k)$ for some c on the line segment connecting a_k to k. Then $c \in Q_k$ so $|\Phi'(c) - \Phi'(a_k)| < \epsilon$. Thus, $|\Phi(x)_j - \Phi(a_k)_j - \Phi'(a_k)_j(x - a_k)_j| < \epsilon|x - a_k|_j$. This then implies $\|\Phi(x) - \Phi(a_k) - \Phi'(a_k)(x - a_k)\| < \epsilon\|x - a_k\|$ since a vector that is larger in each component has larger norm. This local statement translates into the global statement

$$\Phi(a_k) + (1 - \epsilon)\Phi'(a_k)(Q_k - a_k) \subset \Phi(Q_k) \subset \Phi(a_k) + (1 + \epsilon)\Phi'(a_k)(Q_k - a_k)$$

which may be verified by checking it for each $x \in Q_k$, since x must lie in the space between the cubes $\Phi(a_k) + (1 - \epsilon)\Phi'(a_k)(Q_k - a_k)$ and $\Phi(a_k) + (1 + \epsilon)\Phi'(a_k)(Q_k - a_k)$. Now Φ is a bijection so $\Phi(Q_k)$ are almost disjoint, which means their images are also almost disjoint since Φ maps sets of measure zero to sets of measure zero by part (a). Thus, $m(\Phi(\cup Q_k)) = \sum m(\Phi(Q_k))$ and

$$\sum m\left(\Phi(a_k) + (1 - \epsilon)\Phi'(a_k)(Q_k - a_k)\right) \leq \sum m(\Phi(Q_k)) \leq \sum m\left(\Phi(a_k) + (1 + \epsilon)\Phi'(a_k)(Q_k - a_k)\right).$$

Now by Exercise 4 of Chapter 2,

$$\sum m\left(\Phi(a_k) + (1 \pm \epsilon)\Phi'(a_k)(Q_k - a_k)\right) = (1 \pm \epsilon)|\det \Phi'(a_k)| m(Q_k)$$

since Φ' is linear. Thus,

$$(1 - \epsilon)\sum |\det \Phi'(a_k)| m(Q_k) \leq \sum m(\Phi(Q_k)) \leq (1 + \epsilon)\sum |\det \Phi'(a_k)| m(Q_k).$$

But $\sum |\det \Phi'(a_k)| m(Q_k)$ is a Riemann sum for $\int_R |\det \Phi'(x)| dx$, and Riemann integration works because ϕ' is continuous on the compact set R. Thus,

$$(1 - \epsilon)\int_R |\det(\Phi'(x))| dx \leq m(\Phi(R)) \leq (1 + \epsilon)\int_R |\det(\Phi'(x))| dx$$

for all ϵ, so $m(\Phi(R)) = \int_R |\det(\Phi'(x))| dx$.
Now let E be any measurable subset of $\mathcal{O}$. Let $U_n \subset \mathcal{O}$ be open sets containing E with $m(U_n \setminus E) < \frac{1}{n}$. Then $\cap U_n = E \cup N$ where $m(N) = 0$. Then

$$\Phi(E \cup N) = \Phi(\cap U_n) = \cap \Phi(U_n).$$

(In general it is only true that $\Phi(\cap U_n) \subset \cap \Phi(U_n)$, but here we have equality because Φ is bijective.) Since $m(\Phi(N)) = 0$,

$$m(\Phi(E)) \leq m(\Phi(E \cup N)) \leq m(\Phi(E)) + m(\Phi(N)) \Rightarrow m(\Phi(E \cup N)) = m(\Phi(E)).$$

Now $\Phi(U_n)$ are nested sets of finite measure, so

$$m(\cap\Phi(U_n)) = \lim m(\Phi(U_n)) = \lim \int_{U_n} |\det \Phi'| = \int_{\cap U_n} |\Phi'| = \int_{E\cup N} |\Phi'| = \int_E |\Phi'|.$$

Thus, we have (finally) that $m(\Phi(E)) = \int_E |\Phi'|$.

(c) We proved in part (b) that $\int_{\mathcal{O}'} g(y)dy = \int_{\mathcal{O}} g(\Phi(x))|\det \Phi'(x)|dx$ holds if g is a characteristic function χ_E for measurable $E \subset \mathcal{O}$. By linearity, this extends to a measurable simple function $\sum c_j \chi_{E_j}$. Now for f nonnegative, take a sequence $f_n \nearrow f$ of simple functions; then by the Monotone Convergence Theorem,

$$\begin{aligned}
\int_{\mathcal{O}'} f(y)dy &= \int_{\mathcal{O}'} \lim f_n(y)dy \\
&= \lim \int_{\mathcal{O}'} f_n(y)dy \\
&= \lim \int_{\mathcal{O}} f_n(\Phi(x))|\det \Phi'(x)|dx \\
&= \int_{\mathcal{O}} \lim f_n(\Phi(x))|\det \Phi'(x)|dx \\
&= \int_{\mathcal{O}} f(\Phi(x))|\det \Phi'(x)|dx.
\end{aligned}$$

Finally, we can extend to complex integrable f by linearity, since f is a linear combination of four nonnegative integrable functions.

□

Problem 2: Show as a consequence of the previous problem: the measure $d\mu = \frac{dxdy}{y^2}$ in the upper half-plane is preserved by any fractional linear transformation $z \mapsto \frac{az+b}{cz+d}$, where $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ belongs to $SL_2(\mathbb{R})$.

Solution. We first note that such a transformation does in fact map the upper half plane to itself: if $z = x + iy$, then

$$\frac{az+b}{cz+d} = \frac{a(x+iy)+b}{c(x+iy)+d} = \frac{ac(x^2+y^2)+(ad+bc)x+bd}{(cx+d)^2+(cy)^2} + i\frac{(ad-bc)y}{(cx+d)^2+(cy)^2}$$

and since $y > 0$ and $ad - bc = 1$ this has positive imaginary part. Now if we write the map $z \mapsto \frac{az+b}{cz+d}$ in terms of its components as $(x, y) \mapsto (x', y')$, then using the ugly formulas for x' and y' from the above expression, we can compute the even uglier partial derivatives and the Jacobian. However, we can shortcut that by using the fact that the Jacobian is always the square norm of the complex derivative. In case this needs proof, suppose $z \mapsto f(z)$ is a complex differentiable function. Then if $f'(z_0) = \alpha + \beta i$, the linear map $z \mapsto f'(z_0)(z - z_0)$ can be rewritten as

$$(x+iy) \mapsto (\alpha+\beta i)((x-x_0)+(y-y_0)) = (\alpha(x-x_0)-\beta(y-y_0))+i(\beta(x-x_0)+\alpha(y-y_0)),$$

or

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix}$$

which has Jacobian $\alpha^2 + \beta^2 = |f'(z_0)|^2$. Now in our case,

$$f'(z) = \frac{(ad - bc)z}{(cz + d)^2} = \frac{1}{(cz + d)^2}.$$

If we let Φ denote the mapping $z \mapsto z'$ (and equivalently $(x, y) \mapsto (x', y')$), then

$$\begin{aligned}
\mu(\Phi(E)) &= \int_{\Phi(E)} \frac{1}{(y')^2} dx'dy' \\
&= \int_E \frac{1}{(y')^2} |\det \Phi'| dxdy \\
&= \int_E \frac{((cx + d)^2 + (cy)^2)^2}{y^2} |(c(x + iy) + d)^2|^2 dxdy \\
&= \int_E \frac{((cx + d)^2 + (cy)^2)^2}{y^2} ((cx + d)^2 + (cy)^2)^2 dxdy \\
&= \int_E \frac{1}{y^2} dxdy \\
&= \mu(E).
\end{aligned}$$

□

Problem 3: Let S be a hypersurface in $\mathbb{R}^d = \mathbb{R}^{d-1} \times \mathbb{R}$, given by

$$S = \{(x, y) \in \mathbb{R}^{d-1} \times \mathbb{R} : y = F(x)\},$$

with F a C^1 function defined on an open set $\Omega \subset \mathbb{R}^{d-1}$. For each subset $E \subset \Omega$ we write $\hat{E}$ for the corresponding subset of S given by $\hat{E} = \{(x, F(x) : x \in E\}$. We note that the Borel sets of S can be defined in terms of the metric on S (which is the restriction of the Euclidean metric on $\mathbb{R}^d$). Thus if E is a Borel set in Ω, then $\hat{E}$ is a Borel subset of S.

(a) Let μ be the Borel measure on S given by

$$\mu(\hat{E}) = \int_E \sqrt{1 + |\nabla F|^2} dx.$$

If B is a ball in Ω, let $\hat{B}^\delta = \{(x, y) \in \mathbb{R}^d, d((x, y), \hat{B}) < \delta\}$. Show that

$$\mu(\hat{B}) = \lim_{\delta \to 0} \frac{1}{2\delta} m((\hat{B})^\delta),$$

where m denotes the d-dimensional Lebesgue measure. This result is analogous to Theorem 4.4 in Chapter 3.

(b) One may apply (a) to the case when S is the (upper) half of the unit sphere in $\mathbb{R}^d$, given by $y = F(x)$, $F(x) = (1 - |x|^2)^{1/2}$, $|x| < 1$, $x \in \mathbb{R}^{d-1}$. Show that in this case $d\mu = d\sigma$, the measure on the sphere arising in the polar coordinate formula in Section 3.2.

(c) The above conclusion allows one to write an explicit formula for $d\sigma$ in terms of spherical coordinates. Take, for example, the case $d = 3$, and write $y = \cos\theta$, $x = (x_1, x_2) = (\sin\theta\cos\phi, \sin\theta\sin\phi)$ with $0 \le \theta < \frac{\pi}{2}$, $0 \le \phi < 2\pi$. Then according to (a) and (b) the element of area $d\sigma$ equals $(1 - |x|^2)^{-1/2} dx$. Use the change of variable theorem in Problem 1 to deduce that in this case $d\sigma = \sin\theta d\theta d\phi$. This may be generalized

to d dimensions, $d \geq 2$, to obtain the formulas in Section 2.4 of the appendix in Book I.

Solution.

(a) We proceed in the steps outlined in Prof. Garnett's hint:

- Since B is compact and Ω^c is closed, the distance between them is greater than zero, so we can choose $\delta < d(B, \Omega^c)$. Let $V_\delta = \{x : d(X,B) < \delta\} \subset \Omega$. For each $x \in V_\delta$ define $I_\delta(x) = \{y \in \mathbb{R} : (x,y) \in \hat{B}^\delta\}$ and $h(x,\delta) = m(I_\delta(x))$.
- Note that $\hat{B}^\delta \subset V_\delta \times \mathbb{R}$ since for $(x,y) \in \mathbb{R}^{d-1} \times \mathbb{R}$, $d((x,y),\hat{B}) \geq d(x,B)$. By Tonelli's theorem,

$$m(\hat{B}^\delta) = \int_{(x,y)\in V_\delta\times\mathbb{R}} \chi_{I_\delta(x)}(y) = \int_{V_\delta}\int_{\mathbb{R}} \chi_{I_\delta(x)}(y)dydx = \int_{V_\delta} h(x,\delta)dx.$$

- For $x \in B$, let $M = |\nabla F|$ evaluated at x. Let $\vec{v}$ be the unit vector in the direction of ∇F. Then $\nabla F \cdot \vec{v} = M$ at x. By the continuity of ∇F, for any $\epsilon > 0$ we may choose $\delta_0 > 0$ such that $|\nabla F| < M + \epsilon$ and $\nabla F \cdot \vec{v} > M - \epsilon$ at all points within δ_0 of x. Suppose $\delta < \delta_0$. Then for $y \in \mathbb{R}$, if $F(x) \leq y \leq \delta\sqrt{1+(M-\epsilon)^2}$, consider $F(x+t\vec{v})$ for $0 \leq t \leq \frac{\delta(M-\epsilon)}{1+(M-\epsilon)^2}$. By the construction of δ, $F(x) + (M-\epsilon)t < F(x+t\vec{v}) < F(x) + (M+\epsilon)t$. Hence, because $\frac{y(M-\epsilon)}{1+(M-\epsilon)^2} < \frac{\delta(M-\epsilon)}{\sqrt{1+(M-\epsilon)^2}} < \delta$,

$$F\left(x + \frac{y(M-\epsilon)}{1+(M-\epsilon)^2}\vec{v}\right) > F(x) + \frac{y(M-\epsilon)^2}{1+(M-\epsilon)^2}.$$

By the intermediate value theorem, there is some $t_0 < \frac{y(M-\epsilon)}{1+(M-\epsilon)^2}$ at which $F(x+t_0\vec{v}) = F(x) + \frac{y(M-\epsilon)^2}{1+(M-\epsilon)^2}$. Let $x_0 = x + t_0\vec{v}$. Then $(x_0, F(x_0)) \in \hat{B}$; moreover, the distance squared from (x,y) to $(x_0, F(x_0))$ is

$$\begin{aligned} t_0^2 + \left(y - \frac{y(M-\epsilon)^2}{1+(M-\epsilon)^2}\right)^2 &= t_0^2 + \left(\frac{y^2}{1+(M-\epsilon)^2}\right)^2 \\ &\leq \left(\frac{y(M-\epsilon)}{1+(M-\epsilon)^2}\right)^2 + \left(\frac{y^2}{1+(M-\epsilon)^2}\right)^2 \\ &= \frac{y^2}{1+(M-\epsilon)^2} < \delta \end{aligned}$$

so $y \in I_{x,\delta}$. On the other hand, if $y > F(x) + \delta\sqrt{1+(M+\epsilon)^2}$, suppose $(x,y) \in I_{x,\delta}$, so there is some x' with $\text{dist}((x,y),(x',F(x'))) < \delta$. Clearly this implies $|x'-x| < \delta$; suppose $|x'-x| = t$. Then because $|\nabla F| < M+\epsilon$ between x and x', $F(x') < F(x) + (M+\epsilon)t$. Then the distance squared from (x,y) to $(x', F(x'))$ is

$$t^2 + (y - F(x'))^2 \geq t^2 + (y-(M+\epsilon)t)^2 = (1+(M+\epsilon)^2)t^2 - 2y(M+\epsilon)t + y^2.$$

This is a quadratic polynomial in t; the minimum occurs when the derivative is zero, i.e. when

$$2t(1+(M+\epsilon)^2) = 2y(M+\epsilon) \Rightarrow t = \frac{y(M+\epsilon)}{1+(M+\epsilon)^2}.$$

At this point, the value of the quadratic is

$$(1+(M+\epsilon)^2)\left(\frac{y(M+\epsilon)}{1+(M+\epsilon)^2}\right)^2 - 2y\frac{y(M+\epsilon)}{1+(M+\epsilon)^2} + y^2 = \frac{y^2}{1+(M+\epsilon)^2} > \delta,$$

so in fact there is no point in $\hat{B}$ within δ of (x,y), a contradiction. Thus, $y \notin I_{x,\delta}$. By symmetry, the same results hold for $y < F(x)$, so $[-\delta\sqrt{1+(M-\epsilon)^2}, \delta\sqrt{1+(M-\epsilon)^2}] \subset I_{x,\delta} \subset [-\delta\sqrt{1+(M+\epsilon)^2}, \delta\sqrt{1+(M+\epsilon)^2}]$. Thus,

$$\sqrt{1+(M-\epsilon)^2} \leq \frac{h(x,\delta)}{2\delta} < \sqrt{1+(M+\epsilon)^2}$$

for $\delta < \delta_0$. This proves that

$$\lim_{\delta\to 0}\frac{h(x,\delta)}{2\delta} = \sqrt{1+|\nabla F|^2}.$$

– Because ∇F is continuous on the compact set B, it attains a maximum M on B. For any $x \in B$, if $|y - F(x)| > (M+1)\delta$ then $y \notin I_{x,\delta}$, because any point $(x', F(x'))$ within δ of (x,y) would have to have $|x'-x| < \delta$, and then the bound on ∇F implies $|F(x') - F(x)| < M\delta \Rightarrow |y - F(x')| > \delta$. Thus, $\frac{h(x,\delta)}{2\delta} < M+1$ for all δ.

– Since we're interested in the limit of small δ, we can restrict our attention to δ below some cutoff value, say a, where $a < d(B, \Omega^c)$. Then $\int_{V_\delta} h(x,\delta) = \int_{V_a} h(x,\delta)$ because $h(x,\delta) = 0$ for $x \in V_a \setminus V_\delta$. This enables us to take all the integrals over the same region. Moreover, since V_a has finite measure, the constant $M+1$ is integrable over V_a. So by the dominated convergence theorem,

$$\begin{aligned}
\lim_{\delta\to 0}\frac{1}{2\delta}m(\hat{B}^\delta) &= \lim_{\delta\to 0}\int_{V_a}\frac{h(x,\delta)}{2\delta}dx \\
&= \int_{V_a}\lim\frac{h(x,\delta)}{2\delta}dx \\
&= \int_B \sqrt{1+|\nabla F|^2}dx \\
&= \mu(\hat{B}).
\end{aligned}$$

(The region of integration becomes B in the limit because $h(x,\delta)$ is eventually 0 for x outside B.)

(b) For two points $\theta, \phi \in S^{d-1}$, let $a(\theta,\phi)$ denote the angular distance between θ and ϕ, i.e. the angle between the radius vectors to θ and ϕ. Similarly, for a point $\theta \in S^{d-1}$ and a set $E \subset S^{d-1}$, define $a(\theta, E) = \inf_{\phi\in E} a(\theta,\phi)$. Now given a ball $B \subset \mathbb{R}^{d-1}$ and corresponding ball

$\hat{B} \subset S^{d-1}$, define $B'_\delta = \{p \in S^{d-1} : a(p, \hat{B}) < \arcsin(\delta)\}$. I claim that

$$\hat{B} \times [1-\delta, 1+\delta] \subset (\hat{B})^\delta \subset B'_\delta \times [1-\delta, 1+\delta].$$

Here the product is in spherical coordinates, of course. The first inclusion is obvious because if $(\alpha, r) \in \hat{B} \times [1-\delta, 1+\delta]$, then $(\alpha, 1) \in \hat{B}$ and is a distance $|1-r| \leq \delta$ away. For the second inclusion, let (α, r) be any point in $\mathbb{R}^d$. f $\alpha \notin B'_\delta$, let $(\theta, 1)$ be any point in $\hat{B}$. The distance from $(\theta, 1)$ to the line through the origin and (α, r) is $\sin(a(\theta, \alpha))$ by elementary trigonometry. By hypothesis this is greater than δ, so the distance from $(\theta, 1)$ to (α, r) is greater than δ. On the other hand, if $r \notin [1-\delta, 1+\delta]$, then *no* point in S^{d-1} is within δ of (α, r). This proves the second inclusion. Now by the spherical coordinates formulas derived in section 3,

$$m(\hat{B} \times [1-\delta, 1+\delta]) = \sigma(\hat{B}) \int_{1-\delta}^{1+\delta} r^{d-1} dr = \frac{(1+\delta)^d - (1-\delta)^d}{d} \sigma(\hat{B}),$$

so

$$\frac{m(\hat{B} \times [1-\delta, 1+\delta])}{2\delta} = \frac{1}{d} \frac{(1+\delta)^d - (1-\delta)^d}{2\delta} \sigma(\hat{B}).$$

Since $\frac{(1+\delta)^d - (1-\delta)^d}{2\delta}$ is a difference quotient for the function $f(x) = x^d$ at $x = 1$, it approaches $\frac{d}{dx} x^d|_{x=1} = d$ as $\delta \to 0$, so $\frac{1}{2\delta} m(\hat{B} \times [1-\delta, 1+\delta]) \to \sigma(\hat{B})$. Similarly,

$$\frac{1}{2\delta} m(B'_\delta \times [1-\delta, 1+\delta]) = \frac{1}{d} \frac{(1+\delta)^d - (1-\delta)^d}{2\delta} \sigma(B'_\delta).$$

As $\delta \to 0$, this approaches $\lim_{\delta \to 0} \sigma(B'_\delta)$ (provided the latter exists, of course). Now since the B'_δ are nested and $\cap B'_\delta = \overline{\hat{B}}$, $\lim \sigma(B'_\delta) = \sigma(\overline{\hat{B}}) = \sigma(\hat{B})$ by the continuity of measures. (It hardly bears pointing out here that $\sigma(B'_\delta)$ are all finite.) By the Sandwich Theorem,

$$\mu(\hat{B}^\delta) = \lim_{\delta \to 0} \frac{1}{2\delta} m(\hat{B}^\delta) = \sigma(B'_\delta).$$

Thus, $\mu = \sigma$ on balls, but since these generate the Borel sets, $\mu = \sigma$.

(c) By the change of variable theorem,

$$dx = \begin{vmatrix} \frac{\partial x_1}{\partial \theta} & \frac{\partial x_1}{\partial \phi} \\ \frac{\partial x_2}{\partial \theta} & \frac{\partial x_2}{\partial \phi} \end{vmatrix} d\theta d\phi = \begin{vmatrix} \cos\theta\cos\phi & -\sin\theta\sin\phi \\ \cos\theta\sin\phi & \sin\theta\cos\phi \end{vmatrix} d\theta d\phi = \sin\theta\cos\theta d\theta d\phi.$$

Note also that

$$\nabla F = \left(\frac{\partial F}{\partial x_1}, \frac{\partial F}{\partial x_2}\right) = \left(\frac{-x_1}{\sqrt{1-x_1^2-x_2^2}}, \frac{-x_2}{\sqrt{1-x_1^2-x_2^2}}\right)$$

so

$$1 + |\nabla F|^2 = 1 + \frac{x_1^2}{1-x_1^2-x_2^2} + \frac{x_2^2}{1-x_1^2-x_2^2} = \frac{1}{1-x_1^2-x_2^2} = \frac{1}{1-\sin^2\theta} = \frac{1}{\cos^2\theta}.$$

Then

$$d\sigma = \sqrt{1+|\nabla F|^2} dx = \frac{1}{\cos\theta} \sin\theta\cos\theta d\theta d\phi = \sin\theta d\theta d\phi.$$

In n dimensions, one may use for spherical coordinates the n angles $\theta_1, \dots, \theta_n$ with

$$\begin{aligned} y &= \cos\theta_1 \\ x_1 &= \sin\theta_1 \cos\theta_2 \\ x_2 &= \sin\theta_1 \sin\theta_2 \cos\theta_3 \\ &\vdots \\ x_{n-1} &= \sin\theta_1 \dots \sin\theta_{n-1} \cos\theta_n \\ x_n &= \sin\theta_1 \dots \sin\theta_n. \end{aligned}$$

Then the Jacobian $\frac{\partial(x_1,\dots,x_n)}{\partial(\theta_1,\dots,\theta_n)}$ is

$$\begin{vmatrix} c_1c_2 & -s_1s_2 & 0 & 0 & \dots & 0 \\ c_1s_2c_3 & s_1c_2c_3 & -s_1s_2s_3 & 0 & \dots & 0 \\ c_1s_2s_3c_4 & s_1c_2s_3c_4 & s_1s_2c_3c_4 & -s_1s_2s_3s_4 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ c_1s_2\dots s_n & s_1c_2s_3\dots s_n & s_1s_2c_3s_4\dots s_n & s_1s_2s_3c_4s_5\dots s_n & \dots & s_1\dots s_{n-1}c_n \end{vmatrix}$$

where $c_i = \cos\theta_i$ and $s_i = \sin\theta_i$. Call this $J_n(\theta_1, \dots, \theta_n)$. Then we can compute these inductively: a cofactor expansion along the first row yields

$$J_n(\theta_1,\dots,\theta_n) = c_1c_2s_1^{n-1}J_{n-1}(\theta_2,\theta_3,\dots,\theta_n) + s_1s_2s_2^{n-1}J_{n-1}(\theta_1,\theta_3,\dots,\theta_n).$$

Using as our initial case $J_2(\theta_1,\theta_2) = c_1s_1$ as computed above, one has by an easy induction that

$$J_n(\theta_1,\dots,\theta_n) = c_1s_1^{n-1}s_2^{n-2}s_3^{n-3}\dots s_{n-1}.$$

Since $1 + |\nabla F|^2 = \frac{1}{\cos^2\theta_1}$ as before, this yields the area element

$$d\sigma = \sin^{n-1}\theta_1 \sin^{n-2}\theta_2 \dots \sin\theta_{n-1} d\theta_1 \dots d\theta_n.$$

□

Note: If the point of this exercise was to calculate the area element on the unit sphere, it seems that a more direct way is to use the change of variables formula to compute the volume element in spherical coordinates ($r^2 \sin\theta dr d\theta d\phi$ in three dimensions), and since the measure of a set E on the unit sphere is defined to be the measure of the corresponding conical segment, we can compute it as

$$\sigma(E) = 3\int_{[0,1]\times E'} r^2 \sin\theta dr d\theta d\phi = \int_{E'} \sin\theta d\theta d\phi,$$

where E' is the region in the $\theta - \phi$ plane corresponding to the spherical region E. This implies $d\sigma = \sin\theta d\theta d\phi$, with similar formulas in higher dimensions.

Problem 6: Consider an automorphism A of $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$, that is, A is a linear isomorphism of $\mathbb{R}^d$ that preserves the lattice $\mathbb{Z}^d$. Note that A can be written as a $d \times d$ matrix whose entries are integers, with $\det A = \pm 1$. Define the mapping $\tau : \mathbb{T}^d \to \mathbb{T}^d$ by $\tau(x) = A(x)$.

(a) Observe that τ is a measure-preserving isomorphism of $\mathbb{T}^d$.

(b) Show that τ is ergodic (in fact, mixing) if and only if A has no eigenvalues of the form $e^{2\pi ip/q}$, where p and q are integers.
(c) Note that τ is never uniquely ergodic. (Hint.)

Solution.
(a) Duly noted.
OK, I guess I'm supposed to prove it. :-) Since A is a linear isomorphism, A^{-1} is as well. Let $E \subset \mathbb{T}^d$ be measurable, and let $\tilde{E}$ be the corresponding subset of the unit cube in $\mathbb{R}^d$. Then

$$\mu(A^{-1}E) = m(A^{-1}(\tilde{E})) = |\det A^{-1}| m(\tilde{E}) = m(\tilde{E}) = \mu(E)$$

where μ is the measure on $\mathbb{T}^d$ induced by the Lebesgue measure m on $\mathbb{R}^d$.
(b) Suppose that A has an eigenvector of the form $e^{2\pi ip/q}$; then A^T does as well, since it has the same characteristic polynomial. Then $(A^T)^q$ has 1 as an eigenvector. Since $A^T - I$ has all rational entries, there is a (nonzero) eigenvector in $\mathbb{Q}^d$, and hence, by scaling, in $\mathbb{Z}^d$. Let $n \in \mathbb{Z}^d$ with $(A^T)^q n = n$. Consider the function $f(x) = e^{2\pi in\cdot x}$ for $x \in \mathbb{T}^d$. Then since $n\cdot(Ax) = (A^T n)\cdot x$, $T^k f(x) = e^{2\pi i((A^T)^k n)\cdot x}$. The averages of this function are

$$A_m f(x) = \frac{1}{m}\sum_{k=0}^{m-1} e^{2\pi i((A^T)^k n)}.$$

But $T^{k+q}f = T^k f$ for all k, so $A_{jq}f(x) = A_q f(x)$ for any integer j. Since $A_q f(x)$ is not zero (it is a linear combination of exponentials with distinct periods, since we may assume WLOG that q is as small as possible), the averages do not converge a.e. to $\int_{\mathbb{T}^d} f(x)dx = 0$. Hence τ cannot be ergodic.
On the other hand, suppose A has no eigenvector $e^{2\pi ip/q}$. Let $f(x) = e^{2\pi in\cdot x}$ and $g(x) = e^{2\pi im\cdot x}$ for any $m, n \in \mathbb{Z}^d$. Then

$$\langle T^k f, g\rangle = \int_{\mathbb{T}^d} e^{2\pi i((A^T)^k n - m)\cdot x} dx.$$

If $m = n = 0$ the integrand is 1 and the integral is $1 = \langle f, g\rangle$ for all k. If m and n are not both zero, then $(A^T)^k n - m)$ is eventually nonzero; if not, there would be values k_1 and k_2 at which it were zero, but then $(A^T)^{k_1} n = (A^T)^{k_2} n$ so $(A^T)^{k_1} n$ is an eigenvector of $(A^T)^{k_2-k_1}$ with eigenvalue 1. (Since A is invertible, this eigenvector is nonzero). This then implies that A has an eigenvector which is a $(k_2 - k_1)$th root of unity, a contradiction. Thus, $(A^T)^k n - m$ is eventually nonzero, so $\langle T^k f, g\rangle$ is eventually equal to $0 = \langle f, g\rangle$. Hence T is mixing.
As a side note, the fact that (1 is an eigenvalue of A^k) implies (some kth root of 1 is an eigenvalue of A) is not completely trivial. The converse is trivial, of course, but this direction is not if A is not diagonalizable (at least not for any reason that I've found). It follows, however, from the Jordan canonical form. If $J_m(\lambda)$ is a Jordan block of size m with λ on the diagonal, then $J_m(\lambda)^k$ is triangular with λ^k on the diagonal, so it has λ^k as a k-fold eigenvalue. This implies that the algebraic multiplicity of λ^k in A^k is the same as the algebraic multiplicity of λ

in A; since A and A^k have the same dimension, A^k can have no other eigenvalues besides kth powers of eigenvalues of A (since the sum of the multiplicities of the λ_i^k is already equal to the dimension of A^k).

(c) If ν is the Dirac measure, i.e. $\nu(E) = 1$ if $0 \in E$ and 0 otherwise, then $\nu(\tau^{-1}(E)) = \nu(E)$ because the linearity of τ guarantees $0 \in \tau^{-1}(E) \Rightarrow 0 \in E$, and the invertibility of τ guarantees $0 \in E \Rightarrow 0 \in \tau^{-1}(E)$.

□

Problem 8: Let $X = [0, 1)$, $\tau(x) = \langle 1/x \rangle$, $x \neq 0$, $\tau(0) = 0$. Here $\langle x \rangle$ denotes the fractional part of x. With the measure $d\mu = \frac{1}{\log 2} \frac{dx}{1+x}$, we have of course $\mu(X) = 1$.

Show that τ is a measure-preserving transformation.

Solution. Let $(a, b) \subset [0, 1)$. Then

$$\langle \frac{1}{x} \rangle \in (a, b) \Leftrightarrow \frac{1}{x} \in \bigcup_{n=1}^{\infty} (n + a, n + b) \Leftrightarrow x \in \bigcup_{n=1}^{\infty} \left(\frac{1}{n + b}, \frac{1}{n + a} \right)$$

so

$$\begin{aligned}
\mu(\tau^{-1}((a, b))) &= \sum_{n=1}^{\infty} \mu\left(\left(\frac{1}{n + b}, \frac{1}{n + a}\right)\right) \\
&= \sum_{n=1}^{\infty} \frac{1}{\log 2} \int_{1/(n+b)}^{1/(n+a)} \frac{dx}{1 + x} \\
&= \frac{1}{\log 2} \sum_{n=1}^{\infty} \log \left(\frac{1 + \frac{1}{n+a}}{1 + \frac{1}{n+b}} \right) \\
&= \frac{1}{\log 2} \log \prod_{n=1}^{\infty} \frac{(n + a + 1)(n + b)}{(n + a)(n + b + 1)}.
\end{aligned}$$

But

$$\prod_{n=1}^{\infty} \frac{(n + a + 1)(n + b)}{(n + a)(n + b + 1)} = \frac{1 + b}{1 + a}$$

because the product telescopes; all terms cancel except $1 + b$ on the top and $1 + a$ on the bottom. Hence

$$\mu(\tau^{-1}((a, b)) = \frac{1}{\log 2} \log \left(\frac{1 + b}{1 + a} \right) = \mu((a, b)).$$

Since τ is measure-preserving on intervals and these generate the Borel sets, it is measure-preserving. □

Note: By following the hint and telescoping a sum rather than a product, it is possible to prove τ is measure-preserving for all Borel sets directly rather than proving it for intervals and then passing to all Borel sets. Let $E \subset [0, 1)$ be Borel, let $E + k$ denote the translates of E, and $1/(E + k) = \{1/(x + k) : x \in E\}$. Since

$$\frac{1}{1 + x} = \sum_{k=1}^{\infty} \frac{1}{k + x} - \frac{1}{k + 1 + x} = \sum_{k=1}^{\infty} \frac{1}{(k + x)(k + 1 + x)},$$

it follows that

$$\begin{aligned}\mu(E) &= \frac{1}{\log 2}\int_E \frac{1}{1+x}dx \\ &= \frac{1}{\log 2}\int_E \sum_{k=1}^{\infty}\frac{1}{(x+k)(x+k+1)}dx \\ &= \frac{1}{\log 2}\sum_{k=1}^{\infty}\int_E \frac{1}{(x+k)(x+k+1)}dx \\ &= \frac{1}{\log 2}\sum_{k=1}^{\infty}\int_{E+k} \frac{1}{x(x+1)}dx\end{aligned}$$

where we have used the monotone convergence theorem to interchange the sum and the integral. Now if we make the change of variable $x = \frac{1}{y}$, then it turns out that $\frac{dx}{x(x+1)} = \frac{dy}{1+y}$, so this equals

$$\begin{aligned}\frac{1}{\log 2}\sum_{k=1}^{\infty}\int_{1/(E+k)} \frac{1}{y+1}dy &= \sum_{k=1}^{\infty}\mu\left(\frac{1}{E+k}\right) \\ &= \mu\left(\bigcup_{k=1}^{\infty}\frac{1}{E+k}\right) \\ &= \mu(\tau^{-1}(E)).\end{aligned}$$

Chapter 7.5, Page 380

Exercise 2: Suppose E_1 and E_2 are two compact subsets of $\mathbb{R}^d$ such that $E_1 \cap E_2$ contains at most one point. Show directly from the definition of the exterior measure that if $0 < \alpha \leq d$, and $E = E_1 \cap E_2$, then

$$m_\alpha^*(E) = m_\alpha^*(E_1) + m_\alpha^*(E_2).$$

Solution. If $E_{\cap}E_2 = \emptyset$ then $d(E_1, E_2) > 0$ because both are compact, so for $\delta < d(E_1, E_2)$, every δ-cover of $E_1 \cup E_2$ is a disjoint union of a δ-cover of E_1 and a δ-cover of E_2. This implies

$$H_\alpha^\delta(E) = H_\alpha^\delta(E_1) + H_\alpha^\delta(E_2)$$

and taking the limit as $\delta \to 0$ yields $m_\alpha^*(E) = m_\alpha^*(E_1) + m_\alpha^*(E_2)$.

Now suppose $E_1 \cap E_2 = \{z\}$. Let $F_1^\delta = E_1 \setminus B_\delta(z)$ and $F_2^\delta = E_2 \setminus B_\delta(z)$. For any δ-cover $\{F_j\}$ of E, let $\{A_i\}$ be the collection of those F_j which intersect F_1^δ, and $\{B_k\}$ the collection of those that intersect F_2^δ. Note that these collections are disjoint because $d(F_1^\delta, F_2^\delta) \geq \delta$. Then $\{A_i\} \cup \{B_\delta(z)\}$ is a δ-cover for E_1, and $\{B_k\} \cup \{B_\delta(z)\}$ is a δ-cover for E_2. Thus

$$H_\alpha^\delta(E_1) + H_\alpha^\delta(E_2) \leq H_\alpha^\delta(E) + 2\delta^\alpha.$$

Taking limits as $\delta \to 0$, we have $m_\alpha^*(E) \geq m_\alpha^*(E_1) + m_\alpha^*(E_2)$. The reverse inequality always holds, of course, because m_α^* is an outer measure. $\square$

Exercise 3: Prove that if $f : [0,1] \to \mathbb{R}$ satisfies a Lipschitz condition of exponent $\gamma > 1$, then f is a constant.

Solution. Suppose $|f(x)-f(y)| \leq M|x-y|^\gamma$ with $\gamma > 1$. Let $0 \leq x < y \leq 1$ and let $h = y - x$. Then for any integer n,

$$\begin{aligned}
|f(x) - f(y)| &= \left| \sum_{j=0}^{n-1} f\left(x + \frac{j+1}{n}h\right) - f\left(x + \frac{j}{n}h\right) \right| \\
&\leq \sum_{j=0}^{n-1} \left| f\left(x + \frac{j+1}{n}h\right) - f\left(x + \frac{j}{n}h\right) \right| \\
&\leq \sum_{j=0}^{n-1} M \left(\frac{h}{n}\right)^\gamma \\
&= Mh^\gamma n^{1-\gamma}.
\end{aligned}$$

This is true for all n, and the bound approaches 0 as $n \to \infty$, so $f(x) = f(y)$. □

Exercise 4: Suppose $f : [0,1] \to [0,1] \times [0,1]$ is surjective and satisfies a Lipschitz condition

$$|f(x) - f(y)| \leq C|x - y|^\gamma.$$

Prove that $\gamma \leq 1/2$ directly, without using Theorem 2.2.

Solution. Suppose $\gamma > \frac{1}{2}$. By constructing a lattice of spacing $\frac{1}{n}$, we can find n^2 points in $[0,1]^2$ with the property that any two are at least $\frac{1}{n}$ apart. Call these points y_k. For each y_k, let $x_k \in [0,1]$ be any point in the preimage. Then for any $k \neq j$, $|x_j - x_k| \geq \left(\frac{|y_j - y_k|}{C}\right)^{1/\gamma} \geq \frac{1}{(Cn)^{1/\gamma}}$. But for n sufficiently large, $(Cn)^{1/\gamma} < n^2$, so the n^2 points x_k must all be farther than $\frac{1}{n^2}$ from each other. This is manifestly impossible; by the Pigeonhole Principle, some interval $[\frac{j}{n^2}, \frac{j+1}{n^2}]$ for $j = 0, \ldots, n^2 - 1$ must contain two of the points. □

Exercise 5: Let $f(x) = x^k$ be defined on $\mathbb{R}$, where k is a positive integer, and let E be a Borel subset of $\mathbb{R}$.

(a) Show that if $m_\alpha(E) = 0$ for some α, then $m_\alpha(f(E)) = 0$.
(b) Prove that $\dim(E) = \dim f(E)$.

Solution.
(a) Since

$$f(E) = f\left(\bigcup_{n=-\infty}^{\infty} E \cap [n, n+1]\right) = \bigcup_{n=-\infty}^{\infty} f(E \cap [n, n+1]),$$

it is sufficient to show $m_\alpha(f(E \cap [n, n+1])) = 0$. But f is Lipschitz on $[n, n+1]$ because

$$|f(x) - f(y)| = |x^k - y^k| = |x - y||x^{k-1} + x^{k-2}y + \cdots + xy^{k-2} + y^{k-1}|$$

and the second term is continuous on the compact set $[n, n+1]$, hence bounded. By Lemma 2.2, this implies $m_\alpha(f(E \cap [n, n+1])) = 0$.

(b) Let $\alpha > \dim E$. Then $m_\alpha(E) = 0 \Rightarrow m_\alpha(f(E)) = 0$ so $\alpha \geq \dim f(E)$. Hence $\dim f(E) \leq \dim E$. To show the reverse, it suffices to show that $m_\alpha(f(E)) = 0$ implies $m_\alpha(E) = 0$, since then we can apply exactly the same logic with E and $f(E)$ interchanged. Let

$$g(x) = \begin{cases} (-x)^{1/k} & x < 0 \text{ and } k \text{ even} \\ x^{1/k} & \text{else.} \end{cases}$$

Note that $g(f(x)) = \pm x$, so $E \subset g(f(E)) \cup -g(f(E))$. Thus, it will suffice to prove that g is "σ-Lipschitz". Since

$$\mathbb{R} = \bigcup_{n=1}^{\infty} [-n-1, -n] \cup \bigcup_{n=1}^{\infty} [n, n+1] \cup \bigcup_{n=1}^{\infty} \left[-\frac{1}{n}, -\frac{1}{n+1}\right] \cup \bigcup_{n=1}^{\infty} \left[\frac{1}{n+1}, \frac{1}{n}\right] \cup \{0\},$$

and g is Lipschitz on each of these compact sets (it is C^1 on all but the last), the result follows. (If we let $\{K_n\}$ denote all the sets in the above decomposition, then g is Lipschitz on K_n, so $m_\alpha^*(g(f(E) \cap K_n)) = 0$ by Lemma 2.2, and by countable additivity $m_\alpha^*(g(f(E))) = 0$. Similarly $m_\alpha^*(-g(f(E)) = 0$, so $m_\alpha^*(E) = 0$.)

□

Exercise 6: Let $\{E_k\}$ be a sequence of Borel sets in $\mathbb{R}^d$. Show that if $\dim E_k \leq \alpha$ for some α and all k, then

$$\dim \bigcup_k E_k \leq \alpha.$$

Solution. Suppose to the contrary that $\dim \cup E_k > \alpha$. Choose α' with $\alpha < \alpha' < \dim \cup E_k$. Then $m_{\alpha'}(\cup E_k) = \infty$ because $\alpha' < \dim \cup E_k$. But $m_{\alpha'}(E_k) = 0$ for each k because $\alpha' > \dim E_k$, which implies $m_{\alpha'}(\cup E_k) \leq \sum m_{\alpha'}(E_k) = 0$ by countable subadditivity. This is a contradiction, so $\dim \cup E_k \leq \alpha$. □